



Fault-Diagnosis for Control of Supermarket Refrigeration Systems

MASTER THESIS

Author: Zoltán Márk Pintér

Supervisors: Hans Henrik Niemann, Dimitrios Papageorgiou

External supervisor: Lars Finn Sloth Larsen

DTU Electrical Engineering
Department of Electrical Engineering
August, 2019

Fault-Diagnosis for Control of Supermarket Refrigeration Systems

Report written by:

Zoltán Márk Pintér (s172040)

Advisor(s):

Hans Henrik Niemann,
Dimitrios Papageorgiou,
Lars Finn Sloth Larsen

DTU Elektro
Technical University of Denmark
2800 Kgs. Lyngby
Denmark

elektro@elektro.dtu.dk

Project period: 01.03.2019 - 21.08.2019

ECTS: 30

Education: MSc

Field: Electrical Engineering,
Automation and Robot Technology study line

Class: Public

Edition: 1. edition

Remarks: This report is submitted as partial fulfillment of
the requirements for graduation in the above education
at the Technical University of Denmark.

Copyrights: © Zoltán Márk Pintér, 2019

Acknowledgements

This thesis work was conducted at the Department of Electrical Engineering at DTU. It was created in collaboration with Danfoss A/S, RAC. It indirectly includes several people's knowledge and contribution.

I would like to express my gratitude to my team in Danfoss Nordborg, Prins Jan, Torben Green, Salvatore Piscopiello, Peter Reichwald and Ebbe Stubbe. They taught me thermodynamics, the basis of refrigeration systems and their control, opening the gate of a new exciting world for me.

I have always been surrounded with my family and my friends. I would like to thank my father, Csaba, mother, Hajnalka, my siblings, Boróka, Dorka and Brúnó, that they are there for me. I would like to thank my grandmother and grandfather, and I hope that they see it somewhere, what I have become. I would like to thank Cristian and Daniel for the coffee breaks, and all my friends, who prevented me from having lunch alone and crying.

I would like to thank my love, Liza, that she tolerates me, and gives me beautiful moments.

I am grateful to my supervisors, that they guided me along the path of knowledge and decisions, while they did not force anything on me, letting the thesis become one of the most exciting experiences in my life. Thank Lars Larsen for the stimulating conversations on the industrial challenge, that he introduced several of the concepts about the Danfoss control strategies, and spent his invaluable time with me to create the field data experiments. Thank Henrik Niemann, that he always gave wise advices, and guided me in the right direction, when it was necessary. His experience and rationality saved me from several wasted hours. Thank Dimitrios Papageorgiou, that he supported my special course and this work without exhaustion, giving me the feeling that he is just as excited about it, as I am. I hope that he will like this thesis, because it is his work as well.

List of Figures

3.1	The ph (pressure-enthalpy) diagram of CO ₂ (carbon-dioxide)	7
3.2	The refrigeration cycle. Left side: standard setup. Right side: Setup with receiver. Down: Two-level setup. <i>Notation:</i> $cab \sim \text{cabinet}$, $amb \sim \text{ambient}$, $DQ \sim \text{heat flow rate}$	8
3.3	Back-end of the refrigeration system, with 3-cell gas cooler. <i>Notation:</i> $Dm \sim \text{mass flow rate}$, $amb \sim \text{ambient}$, $DQ \sim \text{heat flow rate}$	10
3.4	Back-end ph-diagram of the refrigeration system	10
3.5	Illustration of gas loop. Normal (left) and faulty operation (right). Blue line: fault on measurement, brown line: refrigerant flows. Thickness denotes increasing mass flow rate	11
3.6	Physical structure of a gas cooler (taken from the documentation of TIL Library) .	14
3.7	Temperature distribution along the cells of the gas cooler, for $n_{\text{cell}} = 10$	15
3.8	Modelling of heat transfer. Physical interactions (up), electrical analogy (left), con- vection model (right)	18
3.9	Modelling of the bypass valve. Physical model (left), electrical analogy (right) . . .	21
3.10	Partial differentials along a pressure-enthalpy grid	23
3.11	Gas cooler cell enthalpies for equal volume cells for a 10-cell simulation	26
3.12	Linear least squares transfer parameters along a pressure-enthalpy grid	27
4.1	Illustration of system dynamics. Warm colour: <i>increase</i> , cold colour: <i>decrease</i> , green colour: <i>no effect</i>	30
4.2	Piecewise exponential of the eigenvalues, and the transpose of the transformation matrix \mathbf{W} to the eigenstates. On the left, the colour coding refers to differential relation; on the right, it refers to linear relation	31
5.1	The Modelica model. Green: CO ₂ . Yellow: gas. Red: heat transfer	35
5.2	Simulated ambient temperature	37
5.3	Simulated fault scenarios	37
5.4	Refrigeration system for the field experiment	38
5.5	Treatment of gas cooler outlet temperature data	39
5.6	Fault for field data	39
6.1	The constrainer function. The dashed line denotes the limit value	43
6.2	Innovation signals (down) and trace of covariance matrix (up) for the simulation signal, with no fault handling and insulation fault	45
7.1	Cross covariance between the ambient temperature and the gas cooler outlet tem- perature, normal operation	49
7.2	Residual statistics for simulation data. Statistics of whiteness (up) and fitting normal distribution (down)	51
7.3	Statistical criteria for structure choice. Left: information criteria. Right: identified poles and zeros for two cross-validation sets, with ellipses indicating 95% confidence interval	53

7.4	Residual statistics before and after whitening. Statistics of whiteness (up) and fitting normal distribution (down)	55
8.1	Visualization of the E and M steps [5]. The content of the log likelihood value originally, and after the steps (up); the visualization of the fitting (down)	60
8.2	Illustration of fault tracking. The larger the mean value is, the more the faulty operation cluster tracks it	60
8.3	Fault handling events with the ambient temperature scenario (simulation)	62
9.1	Residuals and fault detection for simulation, in case of insulation fault	65
9.2	Residuals and fault detection for field data	66
9.3	Receiver operating characteristic curves. For simulation evaluation, insulation fault is considered	67
9.4	Reconstruction of outputs for simulation, when insulation fault is ignored	68
9.5	Reconstruction of outputs for simulation, with insulation fault and fault handling	68
9.6	Fault estimation for simulation. Fault handling is active	69
9.7	Fault and inputs for field data	69
9.8	Coefficient of Performance (COP). Fault switch-on and switch-off are indicated by vertical lines. In case of ignored flat fault, the simulation crashed at ≈ 9000 s	71
9.9	Evaporator outlet enthalpies. Fault switch-on and switch off are indicated with vertical lines	71
9.10	Input signals, simulation data, insulation fault	72
9.11	Estimation results for virtual gas cooler enthalpy, in case of insulation fault. The results of the fault accommodating estimator are denoted by dashed lines	73
9.12	Estimation results for virtual gas cooler air temperature, in case of insulation fault. The results of the fault accommodating estimator are denoted by dashed line	74
9.13	Estimation results against theoretical approximations, in case of insulation fault and fault handling. Periodically enclosed area indicates phase lead or lag	74
9.14	Results of constraints for the enthalpy of the receiver, in case of insulation fault and fault ignoring controller	75
B.1	Controllability staircase form	85
B.2	Observability staircase form	85
C.3	Simulink diagram	86
C.4	Fault design	86
C.5	Ambient temperature	87
C.6	Sensor emulator	87
C.7	Fault handling	88
C.8	Recursive linear least squares	88
C.9	Distributed controller	89
E.10	Photo of the field experiment back-end apparatus	90
F.11	State estimation for field data	90

List of Tables

1	Table of physical notations	x
2	Table of mathematical notations	xi
3	Indices and abbreviations	xii
4	Operators	xiv
7.1	Condition number of covariance matrices of identified parameters	54
9.1	Detection times of detectors. Negative value indicates current false alarm length . .	65
9.2	Missed detection rates, in case of handled insulation fault	66
9.3	False alarm rates, in case of handled insulation fault	66
9.4	Sensitivity indices, in case of handled insulation fault	67
9.5	CPU running time of 1 second simulated estimation time (fault handling is applied)	72
D.1	Table of reference values and boundary conditions of the simulations	89

Contents

Acknowledgements	iii
List of Figures	iv
List of Tables	vi
Contents	vii
Nomenclature	x
1 Introduction	1
2 Literature Review	3
2.1 Fault diagnosis of cooling systems	3
2.2 Expectation Maximization for fault detection	5
3 Physical and Mathematical Model	6
3.1 Description of a refrigeration system	6
3.1.1 The back-end and the front-end	7
3.1.2 Components within the system	8
3.1.3 The gas loop problem	11
3.2 State of knowledge on modelling refrigeration systems	12
3.3 The dynamical model description	12
3.3.1 Differential equations framework	12
3.3.2 The nonlinear system of differential equations, without simplifications	13
3.3.3 Derivation of nonlinear differential equations, and constraints	17
3.3.4 Measurements and constraints	21
3.4 Simplifications of the model	22
3.4.1 Reasons for simplifications	22
3.4.2 Simplifications	23
3.4.3 Simplified nonlinear differential equations	25
3.4.4 Measurements and constraints	26
3.4.5 The look-up tables for property conversions	27
4 Linear System Analysis	28
4.1 Linear time-invariant system at a well-defined linearization point	28
4.1.1 The system matrix	29
4.1.2 Eigenvalue analysis	30
4.1.3 Controllability and observability analysis	31
4.2 Linear time-varying system analysis	31

5 The Plants	33
5.1 Description of Modelica model	33
5.2 Description of field data	37
6 The State Estimator	40
6.1 Discretization	40
6.2 Estimator architecture	40
6.3 The linear Kalman filter	41
6.4 The steady-state filter	42
6.5 Extreme values and soft constraints	43
6.6 Set point selection	44
6.7 Tuning of the estimator	44
6.8 Residuals and detectability	45
7 Parameter Estimation and Residuals	46
7.1 Physical background	46
7.2 Recursive least squares algorithm	47
7.3 Regularization and robustness	48
7.4 Delay calculations	48
7.5 The problem with information density	49
7.6 Strong detectability	50
7.7 Parameter identification for the state estimator	50
7.8 Results	51
7.9 Residual whitening	52
7.9.1 Input signal choice	52
7.9.2 Structure choice criteria	52
7.9.3 Results	54
8 Fault Detection and Operation	56
8.1 CUSUM detector	56
8.2 GLR detector	57
8.3 Expectation Maximization detector	58
8.4 Fault handling	62
9 Results and Evaluation	63
9.1 Fault diagnosis and estimation evaluation	63
9.1.1 Evaluation statistics	63
9.1.2 Results	64
9.1.3 Reconstruction of measurement	67
9.2 Control performance analysis	69
9.2.1 Coefficient of performance	70
9.2.2 Results	70
9.3 State estimator evaluation	72
10 Conclusion and Discussion	76
10.1 Thesis summary	76
10.2 Conclusions	76
10.3 Discussion and future work	77
Bibliography	79

Appendices	83
A Conclusion and Perspectives of the Special Course	84
A.1 Conclusion	84
A.2 Perspectives	84
B Observability and controllability staircase forms	85
C Simulink diagrams	86
D Reference values and boundary conditions in the simulation	89
E Field experiment, back-end apparatus	90
F State estimation for the field experiment data	90

Nomenclature

The notations are presented in Table 1 for the physical quantities, and in Table 2 for the mathematical symbols. Indices and abbreviations are described Table 3 and special operators are listed in Table 4.

Table 1: Table of physical notations

Notation	Description
A	Area
d	Diameter
c_p	Isobaric specific heat
CR	Capacity ratio (Utilization of an actuator)
f	Frequency, or degrees of freedom, or friction
H	Height
h	Specific enthalpy
i	Volume state index (values can be $0 \dots n_{\text{cell}}$, GC or R)
J	Joint matrix
k	Thermal conduction slope
K	Number of components
K_v	Valve constant
l	Length
\dot{m}	Mass flow rate
N	Number of phases
p	Pressure
q	Vapour quality
\dot{Q}	Heat flow rate
R	Resistance
Re	Reynold's number
t	Time
T	Temperature, or time interval
u	Specific energy
U	Heat transfer coefficient
v	Specific volume, or measurement noise
s	Entropy
V	Volume
\dot{V}	Volumetric flow rate
w	Thermal conduction ratio
W_t	Technical work
x	Receiver liquid level (height)
μ	Dynamic viscosity
Ψ	Excitation
ρ	Density

Continued on next page

Table 1 – continued from previous page

Notation	Description
σ	Thermal conduction
σ_0	Natural convection
τ	Time constant (Slowness of an actuator)

Table 2: Table of mathematical notations

Notation	Description
A	System matrix, or polynomial coefficient vector
B	Input matrix, or polynomial coefficient vector
C	Output matrix, or polynomial coefficient vector
\mathcal{C}	Controllability matrix
D	Feedforward matrix, or dimension
\mathbb{E}	Expected value
f	Nonlinear state derivative function, or fault evaluation signal
\mathcal{F}	F-score
g	Decision function, or nonlinear output function
h	Threshold
\mathcal{H}	Hypothesis
K	Number of clusters, or Kalman gain
KL	Kullback-Leibler divergence
\mathcal{L}	Lower bound
M	Window length
\mathcal{M}	Model
n	Number, or white noise
N	Number of data
\mathcal{N}	Normal distribution
\mathcal{O}	Observability matrix
P	State covariance
p	Parameter, or true probability distribution
q	Delay operator, or model probability distribution
Q	Input state covariance matrix
R	Measurement covariance
R_ε	Innovation covariance
s	Schur complement, or log likelihood ratio
S	Cumulative sum
T	Transformation matrix
T_s	Sampling time
u	Input, or linear input
U	Nonlinear input
v	Measurement noise
W	Loss function
x	State, or linear state
X	Nonlinear state
y	Output, or linear output

Continued on next page

Table 2 – continued from previous page

Notation	Description
Y	Nonlinear output
z	Observation
Z	Latent variable, or inverse of cumulative normal distribution
α	Constant
β	Constant
γ	Responsibility
∂	Partial differential operator, or transfer parameter
ε	Residual
ϕ	Regressor
λ	Lambda
μ	Mean
π	Mixing coefficient
σ	Standard deviation
Σ	Covariance of signal
θ	Estimated parameter, or distribution parameter
ξ	Constant

Table 3: Indices and abbreviations

Notation	Description
0	Normal operation, or ambient
1	Faulty operation
A	Air
AIC	Akaike information criterion
ARL	Average run length
ARMA	Autoregressive, moving average
ARMAX	Autoregressive, moving average, exogenous
AUC	Area under the receiver operating characteristic curve
BC	Boundary condition
BIC	Bayes information criterion
BP	Bypass valve
cell	Cell (in gas cooler)
comp	Compressor
cont	Continuous
conv	Convection
C	Cooler, or controllability staircase form
CO2	Carbon-dioxide
COP	Coefficient of performance
CUSUM	Cumulative Sum
disc	Discrete
dist	Disturbance
eig	Eigenvalue
eq	Assuming equilibrium

Continued on next page

Table 3 – continued from previous page

Notation	Description
EM	Expectation Maximization
f	Filtered
fin	Fin (in gas cooler)
F	Freezer
FPE	Final prediction error
FPR	False positive rate
G	Gas
GC	Gas cooler
GLR	Generalized likelihood ratio
GMM	Gaussian mixture model
hyd	Hydraulic
HR	Heat recovery unit
HV	High pressure valve
HVAC	Heating, ventilation and air-conditioning
id	Ideal
in	Inlet
IT	Intermediate temperature
lim	Limit
L	Liquid
LT	Low temperature
LTI	Linear time-invariant
LTV	Linear time-varying
m	Measurement
MA	Moving average
MDR	Mean density ratio
MT	Medium temperature
nom	Nominal
nonlin	Nonlinear
out	Outlet
O	Observability staircase form
pitch	Pitch temperature (in gas cooler)
PCA	Principal Component Analysis
PI	Proportional-integrator
raw	Not entirely filtered, raw value
R	Receiver
RLS	Recursive (linear) least squares
ROC	Receiver operating characteristic
RV	Receiver valve
Q	Disturbance
s	Sampling
ss	Stationary, or steady-state
SVDD	Support vector data description
TPR	True positive rate
UIO	Unknown input observer
V	Valve

Table 4: Operators

Notation	Description
$\dot{\square} = \frac{d\square}{dt}$	Time derivative
$\frac{\partial}{\partial\square}$	Partial differential
$\hat{\square}$	Estimation
$\tilde{\square}$	Error between true and estimated value
$ \square $	Conditional on \square (probability theory)
\square	Vector or matrix
\parallel	Divergence of distributions (information theory)
$\square_1(\square_2, \square_3)_{ss}$	Steady-state conversion from \square_2 and \square_3 to \square_1

Note that the dependencies of partial differentials are not expressed in the documentation. However, every partial differential depends on two properties - which describe the equilibrium entirely -, for example pressure and enthalpy. This means that $\frac{\partial z}{\partial x}|_y(x, y)$ will be denoted as $\frac{\partial z}{\partial x}|_y$.

Chapter 1

Introduction

In this thesis work, gas loop is to be solved, which is a common problem in the refrigeration industry. The problem appears in supermarket refrigeration systems, when a sensor produces wrong measurements. This, if no action is taken, results in decreasing coefficient of performance (COP, the ratio of the energy transferred and the work invested) of the system, and the cabinet temperatures can no longer be controlled.

Although the COP-wise optimal control strategy (see [54] and [50]) for the system exists, when the gas loop occurs, the conventional strategy against it solves only the effects by temporarily significantly decreasing the COP value. After a while, the system ends up in the original stage, which results in a repetitive behaviour between recurrence of the effects and the counteraction of them.

The goal of the thesis is to detect and accommodate the fault early enough, so that gas loop is not even entered. This is to be done along a broader range of operation points, but it should be small enough for deeper analysis of certain behaviours.

The scope of the thesis is establishing a model-based method of state estimation, finding a strategy to detect the faults and reconstructing the measurements, doing fault accommodation, and doing validation on a wide domain of operation scenarios from 10 to 30 °, including an open loop field experiment. Open loop refers to a scenario, when the fault diagnosis happens offline, and the reconstructed measurement is not used for fault accommodation.

The strategy chosen for fault diagnosis is separating the residual generation from the fault detection problem. The residual is to be generated by a partially model based method, so that more robustness is guaranteed for extensions to new models, than in case of data driven methods. More robustness is expected, since with the law of physics, the most important traits shall be captured. Small range of offline data should be eligible, since generating useful data of refrigeration systems requires the change of seasons. The algorithms should be computationally cheap and run online, so that they can be applied in the industry. The fault accommodation is supposed to be solved by estimating the true value for the sensor, and replacing the faulty measurement value. A specific fault handling scenario is to be set-up, so that the problem can be conveniently remedied.

There will be two ways of evaluating the results. One of them is showing if the coefficient of performance, or the cabinet temperature is disturbed by the fault, and if the problem is solved by fault handling. This is done in a closed loop simulation. The second way is to investigate if the fault is estimated well in an open loop field data experiment (the fault is not considered by the controller there).

The content of the thesis is the following. Chapter 2 provides an overview of the existing fault diagnosis strategies. Chapter 3 discusses the nonlinear mathematical model. The problems of the

first model are described, and a simplified model is presented. Chapter 4 describes the linear analysis on the linearized model. Chapter 5 presents the physical plants. The first one is a simulation model programmed in Modelica language. The second one is an existing refrigeration system at the company, for which the thesis is done, Danfoss. Chapter 6 introduces a two-fold estimator to estimate the states of the model. In Chapter 7, the Recursive Least Squares (RLS) algorithm is proposed as an alternative residual generation method, which is partially data-driven. The residual is whitened, to be able to use more fault detectors. Chapter 8 details the design of fault detectors, which are based on well-known Cumulative Sum (CUSUM) and Generalized Likelihood Ratio (GLR) algorithms. An additional fault diagnosis method, Expectation Maximization (EM), is converted to its online version with fault tracking, to make detection on the coloured residuals. Chapter 9 assesses the fault-tolerance of the system, based on criteria that relate to reconstruction of measurements, control efficiency, and state estimation quality. Finally, Chapter 10, summarizes the findings of the thesis work, and discusses applicability and limitations of the methods in real life.

Chapter 2

Literature Review

In this chapter, the contexts of the fault diagnosis methods used on cooling systems, and the fault diagnosis applications of the Expectation Maximization (EM) algorithm are studied.

2.1 Fault diagnosis of cooling systems

The literature mainly concentrates on the fault diagnosis, isolation, and sometimes the estimation of faults of heat, ventilation and air conditioning (HVAC) and evaporator systems. Faults are ranging from sensor or actuator faults to leakages. There are two basic approaches: model based and data driven. Often, these approaches are mixed.

Model based approaches work better on a well-known system, or simulation model, but become fragile, when extending on new systems. In [32], a virtual temperature sensor was created for an HVAC system. Steady state thermodynamical and empirical equations were used with linear regression, to find the mass flow rate. Six sensors were used for heating and four for cooling, this way faults were isolated for a wide range of operating conditions. In [31] HVAC mass flow rates were estimated by empirical and ANSI/ARI 540-1999 equations. Linear regression was applied again. The equations used were customized for the individual components, for example for fixed and variable speed compressors. Bayesian log-likelihood estimation was used to find if there was a fault. The refrigerant flow faults were successfully decoupled from other faults. Linear analytical equations were used to find the location of sensors in the refrigerant cycle in [23]. A test signal at a given frequency was used for configuration. This was active identification, in a range which was not too corrupted by disturbances and noise. Assuming that the frequency signal acts with different magnitudes on the measurements of different signals, GLRT was used to identify sensors. The effect of the residuals at the test frequency was damped with a weight matrix. Lab experiments showed worse results, than the simulation ones.

In the Kalman-filter based fault isolation [63], half empirical-half, analytical chiller model from [37] was applied for modelling. This model was discretized and the residual was normalized by the variance, afterwards, CUSUM was applied. Two isolation methods were determined. Applying a filter for one sensor is better for isolating multiple fault scenarios, while application for all but one sensor was more robust to plant deviations (see [26]). The likelihood of the model was maximized with Multiple Model Adaptive Estimation. The extended nonlinear version worked better, than the linear one. Unknown Input Observer was applied in [64], for the model from [37]. Wide variety of conditions were tested, and it turned out that the UIO did not work well during the night; it was not as successful as the one in [63] for sensor fault isolation, but performed better in parametric fault isolation.

On the contrary to the model based methods, data driven ones are plug and play, as soon as

sufficiently large and rich volume of offline data are provided. Note that they were not always entirely data based, often they use some logic, or steady-state physical model to improve their results. In [33], fault diagnosis was done with neural network and fuzzy logic. Linguistic rules determined a rule base, upon which the ANN was built. Seven different faults were detected and isolated, and sensor malfunctions were identified as well. The method results in linguistic decisive functions. It was empirically shown that the condenser temperatures have bigger range than the evaporator ones. Low charge detection was done with Eureqa software in [51]. Non-critically charged refrigeration systems were investigated. The symbolic regression applied in the software was based on an evolutionary computation method. It aims at finding mathematical expressions for data with a rule base, while minimizing certain error metrics. Fault detection was achieved with energy signals in [58] for HVAC systems. Energy consumption of the actuators in the front-end were investigated. A SARIMA model was identified: composing of ARMA and seasonal ARMA. Defrost cycles were excluded from the timeseries, in order to provide useful results. Segmented linear trends were assumed for modelling, to overcome nonlinearities. Without using sensors, the results were not convincing, but including them increased performance. Decision tree based fault diagnosis was done in [38] for HVAC systems. The same sensors were used as in the virtual sensor article (half-analytical half-experimental models). Support vector regression was used to fit nonlinearities more. CART algorithm created the applied trees. Pruning was carried out against overfitting, using expert knowledge rules and statistical methods. Tests on varying operating conditions but different training conditions gave 69.% accuracy.

In the sensitivity analysis of principal component based (PCA) based fault detection [25], every sensor got their own Q-test threshold. Eight sensor measurements and the energy balance equations of the chiller were used. SCREE test provided the optimal number of principal components. Shapiro-Wilk test showed in the Support Vector Data Description (SVDD) based chiller fault detection [40], that the variables in refrigeration systems were rather non Gaussian. Eight or sixteen variables were used to compare to PCA-PC subspace methods. It outperforms them, since PCA assumes Gaussian distribution, linear process variables, and it was not adapting to several operating conditions. In the PCA-R-SVDD based chiller fault detection paper [39], the following methods were compared: PCA, SVDD, PCA-PC-SVDD and PCA-R-SVDD, where R denotes residual subspace, and PC denotes principal component subspace. For PCA, not only the Q, but the T (so called transformation or score matrix) boundaries were introduced. For PCA-X-SVDD, where X denotes either of the subspaces, SVDD was applied on the subspaces, then the analytical equations of the two methods were merged. There were hyperparameters for SVDD, which needed to be tuned with 10-fold cross validation. The kernel was Gaussian. The disadvantages of SVDD were the following: it cannot take uncertainty into consideration, has higher false alarm rate for training on data with limited operating conditions, and has high computational complexity. Steady state detection from [30] was used, with forgetting factor. For less severe fault levels, PCA-R-SVDD outperforms the other methods. PCA-R-BN method was used in [61], which was introducing Bayesian networks for classification. Since stand-alone Bayesian networks do not have good detection accuracies for low fault levels, PCA-R-BN architecture was tested. It was compared to PCA, PCA-PC, and Bayesian networks. Two levels were used to define probability layers, fault and symptom levels, then the principle of the maximum posterior probability was used. Maximum likelihood was used to find the conditional probability parameters, and expert knowledge was used to find prior probability parameters. Distance rejection model was used to detect faults in [59]. Nine variables were selected for detection, and the methods were tested on a physical plant. In the adaptive PCA method for chiller fault detection [24], HVAC systems were investigated. Q-boundary rule was used to find new trends in the PC subspace. The word adaptive refers to that it was trained in normal operation by removing rows not fitting the normal assumption. This resulted in a less conservative threshold. Steady state detection from [30] was used. In [49], high evaporator temperatures were detected. Logistic regression, decision trees and random forest (ensemble model of decision tree with bagging and subspace sampling) were tested on industrial refrigeration systems. The last one

outperformed the others in F1-score.

In several of the data based method motivated papers, the steady state detection technique from [30] was used. This introduced an online method for calculating variance; then, with well adjusted window size, physical HVAC signals were investigated, and the thresholds for standard deviation were tuned. Different entering and leaving levels were used to show, that in case of fault (refrigerant undercharge), the method was still operating correctly.

In this work, the base motivation is to establish a method, where extensive amount of offline data is not necessary. For this, analytical model is to be set up; however, as it turns out from the literature review, robustness is a weakness of the model based methods. Note, furthermore, that the data-driven methods are often plug and play, and do not separate the challenges of residual creation and fault diagnosis; on the contrary, this separation will be applied here, based on [6]. In this book, several residual generation strategies are considered, for example structural analysis, Kalman filter innovation, or input-output methods. For fault diagnosis, the single and multi-dimensional versions of the CUSUM and the GLR algorithms are considered, deducing several statistics, like Average Run Length, or probability of false alarm or missed detection. Furthermore, fault reconfiguration and accommodation strategies are discussed. In this work, the latter is applied, by replacing the sensor measurement, before the estimator or the controller uses it. The strategies presented in the book have other advantages as well: they can be used as online algorithms, and are computationally cheap.

2.2 Expectation Maximization for fault detection

Chapter 8 introduces a well known maximum likelihood maximizer for latent variable models [5], the Expectation Maximization (EM) method.

In [22] EM computation optimization is done for Gaussian Mixture (GMM) models. Only one parameter is optimized in one iteration. It is shown that the coefficient performance coefficient (CPC) drops from 90% to 85%, but the cost time improves with three orders of magnitude. In [56] EM is used for GMM to set-up a normal operation model of satellite sensor behaviour, then a combination of particle filtering, machine learning and voting techniques are used together with correlation clustering of sensors, for anomaly detection. In [2], either two (detection) or six (isolation) clusters of GMM were set-up and tested for, using current and drift signals of the motors simulating an AUV unit. Both applications were successful. The problem of missing data is considered in [65]. Conventionally missing data are ignored; in this paper, the partial consideration of non-entire data-vectors are considered. Likelihood is maximized with discarding the missing dimensions for the individual data points. The method is using a closed form solution for multiple missing data pattern; in other words, the data is not expected to be missing always at the same location. The results are tested on simulation of a physical problem, missing 50% of the data, and improvement is shown with regard to the conventional methods. EM algorithm is used in [44], as an augmentation of the Kalman filter, for the detection and isolation of fault profiles of an electro-hydraulic system. In the forward Kalman step, the parameters are found, in the backward step, fault evaluation is done. This algorithm is online. The method is used to reconstruct the correct behaviour in case of sensor or actuator faults, showing the occasional limitations for certain perturbation frequencies with cross-power spectral density plots.

Chapter 3

Physical and Mathematical Model

3.1 Description of a refrigeration system

The description of a refrigeration cycle, which is very similar to the one studied in this thesis, as well of its nonlinear simulation can be found in an earlier work of the author, which was a special course at DTU (see the conclusion in Appendix A). Relevant results of it will be restated in this chapter.

The equilibrium points of refrigeration systems are described by pressure-enthalpy (ph) diagrams. For the current refrigerant, carbon-dioxide, it can be seen in Figure 3.1. The figure depicts the most important thermodynamic properties (from now on just properties) of the thermodynamic states. The thick black line is the saturation curve, with liquid phase from the left of it, while there is gas on the the right side. Above it, at high pressures, the supercritical zone is defined, where the gas and liquid phases cannot be differentiated. Under it, there are both phases, and their ratio (the so called vapour quality) is the function of the enthalpy, and the temperature does not change along the horizontal lines. The horizontal lines are the isobaric ones, meaning that the pressure is constant; the red lines are the isothermals, where the temperature is; the green lines are the iso-chors, where the density, or specific volume is; the blue ones are the isentropic, where the entropy is not changing. The specific enthalpy is an energy measure of open systems,

$$h = u + pv, \quad (3.1)$$

where u is the specific energy and p is pressure. Open systems are where the investigated system (bounded by the so called control surface) has energy and matter exchange with its environment. Control volume is the volume, in which the thermodynamical properties are investigated with conservation laws [11]. The definition of entropy is found in [11]. The values are called specific, since they are normalized by mass, hence there is a better way of comparison of energy levels in the refrigeration cycle.

A standard refrigeration cycle (see Figure 3.2, left side) has two pressure levels, since the temperature increases with the pressure. Larger temperature difference with respect to the ambient temperature results in more effective heat transfer for the same mass flow rate in the gas cooler tubes - as the heat transfer increases by the temperature differences. The compressors and the valves separate the pressure levels. Both of them are used to control the pressure at their inlet. On each of the pressure levels, there are heat exchangers, which are called condenser and evaporator. Condensing and evaporation are used to leverage that there is no change in the temperature along the varying enthalpy, ensuring more effective heat transfer. Note, however, that this system works with carbon-dioxide refrigerant, which needs to go over the critical point (see Figure 3.1, where the critical point is the top of the saturation curve) for achieving large temperature difference, and hence good heat transfer (see [12]). In the right side of Figure 3.2, it can also be seen that a buffer

unit called receiver is used, and this introduces a new, intermediate pressure level (see [10] and [9]). This is not only a mass buffer, but it separates the liquid from the gas (as the fluid leaving the high pressure valve, is two-phase). This is due to the fact, that the liquid provides better heat transfer characteristics, and it is easier to pipe. Another notable consideration is that the compressor's lifetime decreases significantly, if liquid instead of gas enters the suction (inlet) side. The receiver pressure is controlled by the receiver valve.

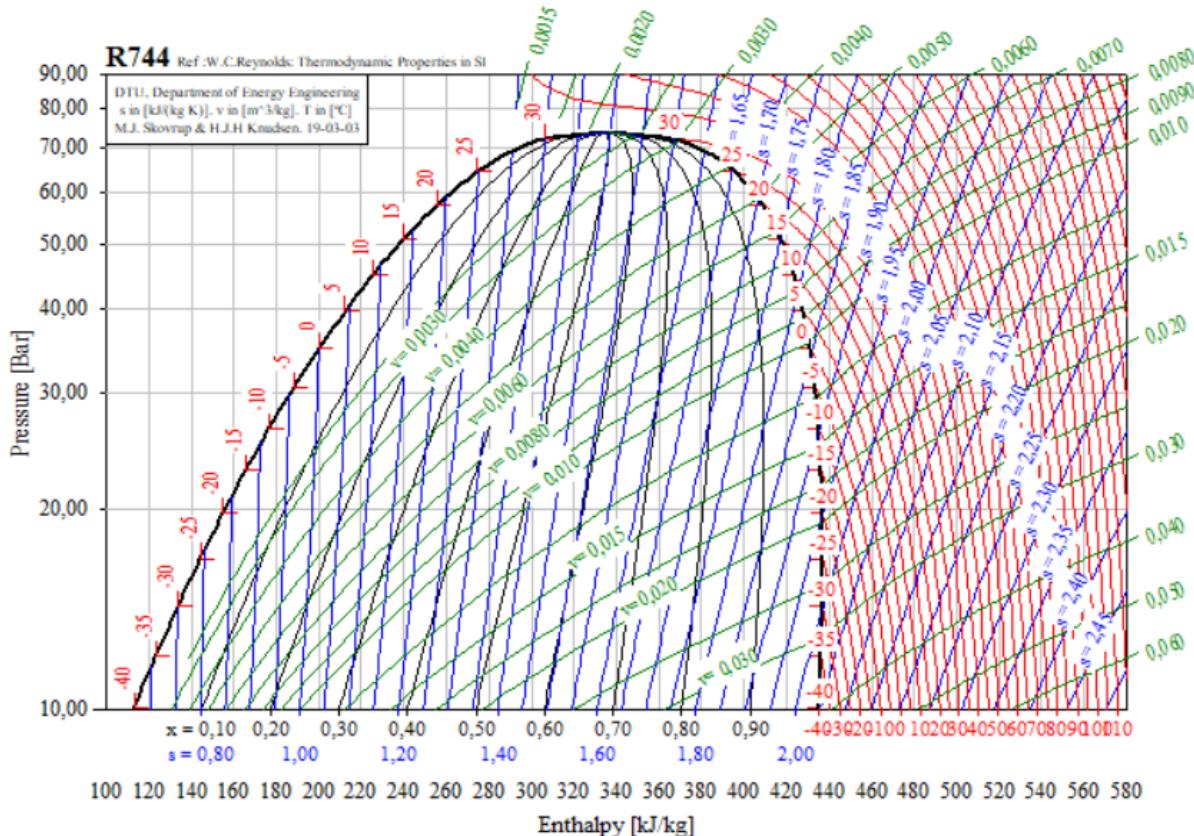


Figure 3.1: The ph (pressure-enthalpy) diagram of CO₂ (carbon-dioxide)

One more extension to this model is introducing the low and middle temperature pressure levels (see Figure 3.2, down). This is again due to the temperature-pressure interdependency, which gives the possibility to control different temperature levels by regulating the pressure levels, like the ones in the cooler and the freezer. The system is called a booster system [17], as the low-temperature (LT) compressor raises the pressure from the lowest level to the medium temperature (MT) level, using the same medium.

3.1.1 The back-end and the front-end

There are several timescales and dynamics appearing in a refrigeration system, which range from sound dynamics through mechanics and fluid dynamics to thermodynamics. Even on the thermodynamic level, the high pressure and receiver pressure level dynamics are considered as being independent from the pressure levels of the evaporators. Evaporator control is more distributed and therefore the dynamics are faster, while the gas cooler becomes the slowest element of the system. The evaporators are called 'front-end', since it is in the supermarket customer area, while the 'back-end' is the condenser with the compressors and the receiver, since it is located in a separate room. In the industry, the control of these two main subsystems are independent from each other.

Hence, within this work, only the back-end (high pressure and receiver pressure) is considered, since this one is more affected by the gas loop problem.

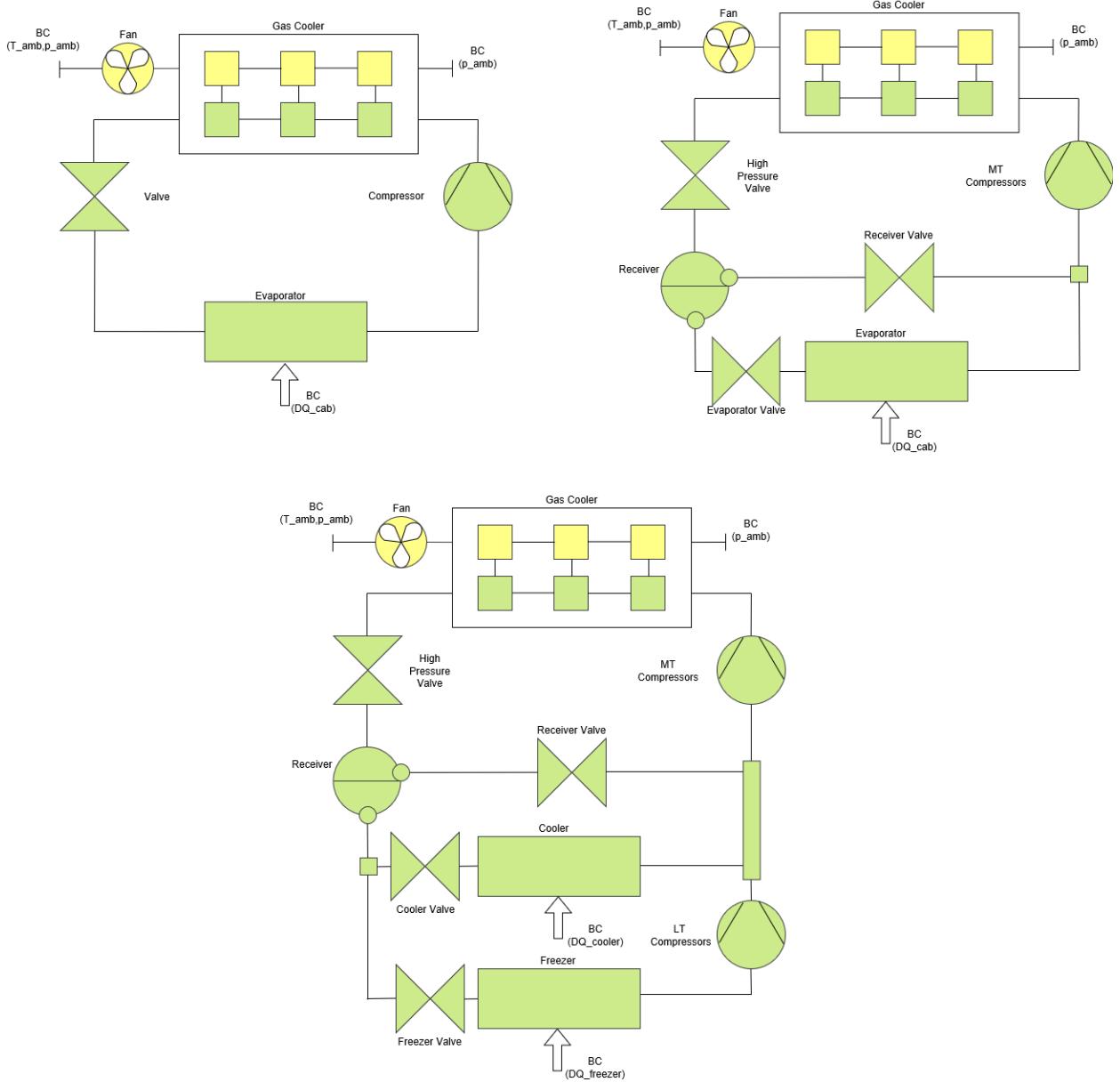


Figure 3.2: The refrigeration cycle. Left side: standard setup. Right side: Setup with receiver. Down: Two-level setup. *Notation:* $\text{cab} \sim \text{cabinet}$, $\text{amb} \sim \text{ambient}$, $DQ \sim \text{heat flow rate}$

3.1.2 Components within the system

The following components are considered from the beginning of the project, from which some may be neglected later. They are depicted, together with the boundary conditions of the system in Figure 3.3. The corresponding equilibrium points on the ph diagram are shown in Figure 3.4. Note that the red letter combinations refer to later indexing in this work.

- Evaporators carry out heat exchange between the cabinet air volume and the refrigerant flowing through its tubes.
- This heat exchange is regulated by the evaporator valves (their operation itself does not change the energy level in the system). These set the mass flow rate through the evaporator.

Larger mass flow rate means cooler air volume, but less efficiently used refrigerant.

- Compressors are used to set the mass flow rate towards a higher pressure, and for this, they increase the energy in the system. They have suction and discharge sides, and they regulate the former one, to keep a certain pressure level. The system is called booster (see [17]), since the low temperature fluid is first pumped up to the medium temperature pressure level by the low temperature (LT) compressor, then it is mixed with lower enthalpy fluids, and pumped up to the high temperature level by the medium temperature (MT) compressors. This strategy enhances efficiency, since the lower enthalpy the suction side has, the smaller enthalpy excess it provides, and less energy is to be dissipated by the gas cooler.
- The gas cooler dissipates the accumulated energy towards the ambient temperature through heat exchange. Its pressure level reference value is regulated as a function of the ambient temperature, so that the heat flow rate settles at the desired level, providing the best efficiency of the overall system [54].
- The high pressure valve is at the outlet of the gas cooler, and it sets the mass flow rate through it, hence regulating the pressure level. The more the valve is open, the lower the pressure level will be.
- The fans are an extra degree of freedom, to avoid integration of energy in the system. They change the volumetric flow rate of the air through the gas cooler, therefore regulating the heat exchange. The energy integration is avoided, since the ambient air can be modelled with infinite heat capacity.
- The receiver is a buffer unit, which is at the suction side of the evaporator valves and at the discharge side of the high pressure valve. It serves to tolerate the contradicting transients of the controlled physical quantities, and to avoid gas being used for heat exchange. This avoidance increases heat capacity of the refrigerant for the heat exchange. Its energy level can be described by its filling level (level of liquid phase), and in case of proper control, it should always contain two-phase liquid.
- The pressure in the receiver is regulated by the receiver valve. Only the less useful gas is bypassed here to the suction side of the MT compressor.
- If the gas is in excess in the receiver, the IT compressors can be used instead of the receiver valve, pumping up the gas immediately to high pressure. Since starting from a higher pressure level means a smaller enthalpy excess introduced to the system (see the isentropic curves, along which the process happens in Figure 3.1), this leads to a more energy efficient process.
- The bypass valve lets the fluid partly bypass the gas cooler in case of cold weather. This is necessary due to the natural convection, which would overcool the fluid and let liquid into the compressors, reducing their lifetime. Fan usage results in forced convection, and when they are switched off, only natural convection is present.
- The heat recovery unit is a heat exchange before the gas cooler. It saves some energy dissipation by using this energy to heat exchange with the water heating system of the supermarket.

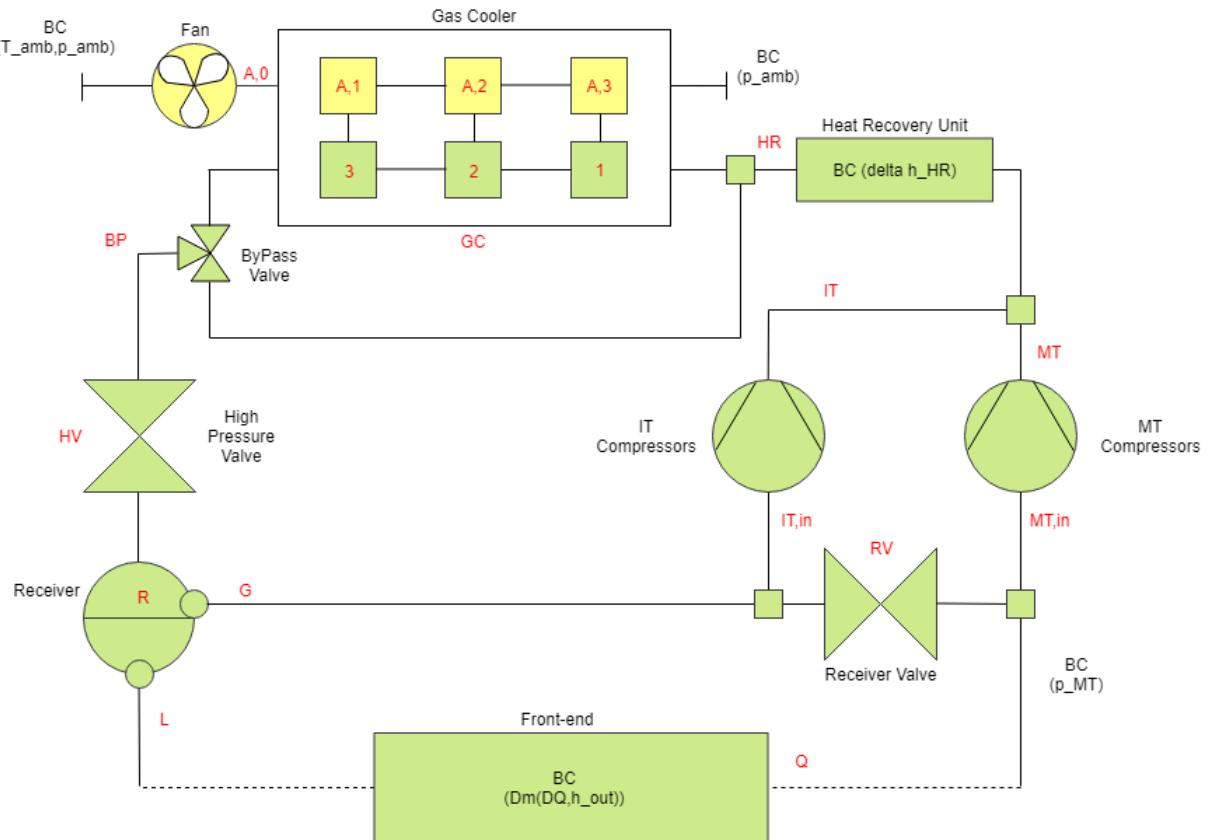


Figure 3.3: Back-end of the refrigeration system, with 3-cell gas cooler. Notation:
 $Dm \sim$ mass flow rate, $amb \sim$ ambient, $DQ \sim$ heat flow rate

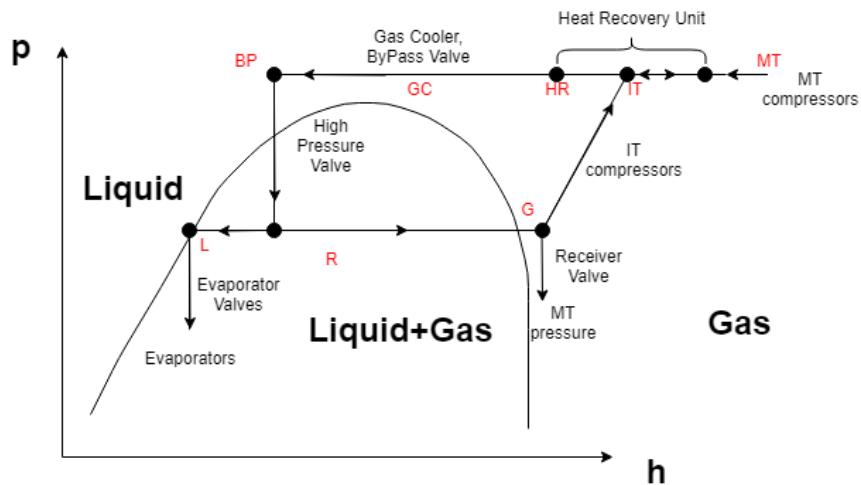


Figure 3.4: Back-end ph-diagram of the refrigeration system

The components can be distinguished into two types, volume components and concentrated components. The volume components let the basic dynamics of thermodynamical processes to occur within them, receiving and passing mass flow rates, which are conditioned at certain enthalpy levels. They have the thermodynamical properties, like density, enthalpy, pressure, temperature, as variables. One example is the receiver. If volume components have a large area with regard to their volume, they might be considered as distributed elements, as the temperature difference, hence all the properties will change along their spatial distribution. These are the heat exchangers.

The concentrated elements are modelled with no volume, receiving pressures from both sides, and enthalpy at their inlet; passing enthalpy at their outlet and opposite sign mass flow rates on both sides. Their variable, if any, can be the mass flow rate or volumetric flow rate.

3.1.3 The gas loop problem

A general fault occurring in supermarket refrigeration systems is a shift in the reading of the temperature sensor at the outlet of the gas cooler. This reading is used for the control of the outlet enthalpy at that location, by regulating the speed of the fans. The shift might be caused by a sensor insulation problem, or an unlocking binding. This way, the ambient temperature appears indirectly on the sensor reading.

The causation is illustrated in Figure 3.5. Line thickness denotes the amount of mass flow rate. The location on the ph diagram, where the sensor shifts, provides a high enthalpy-temperature gradient, resulting in a large enthalpy difference experienced. The controller reduces the forced convection through the fan speed, and the true enthalpy, and hence temperature increases in both the high and the receiver pressure levels. Increasing temperature means increasing pressure, resulting in the IT compressors or the receiver valve working on high capacity, sometimes even saturating. The mass flow rate, instead of being used for heat transfer in the evaporators, is circulating on the two highest pressure levels of the system.

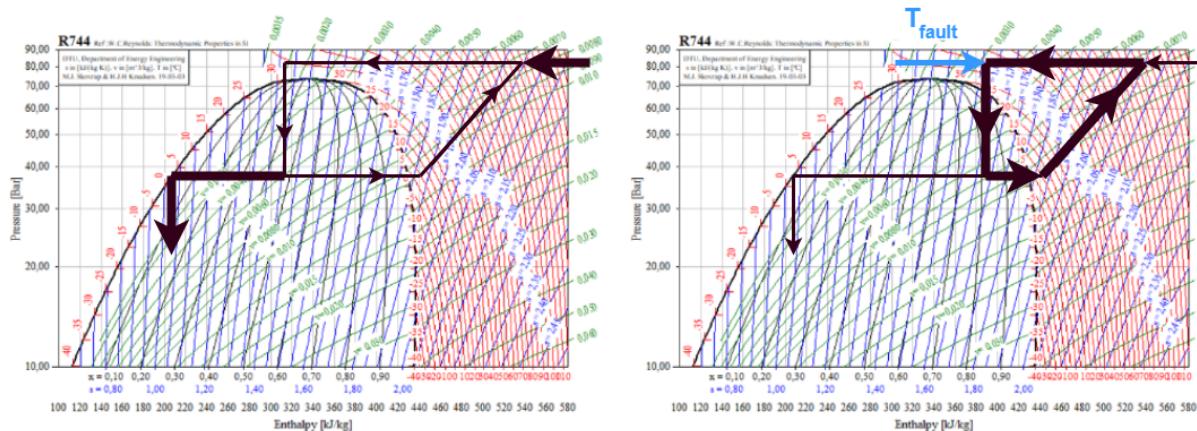


Figure 3.5: Illustration of gas loop. Normal (left) and faulty operation (right). Blue line: fault on measurement, brown line: refrigerant flows. Thickness denotes increasing mass flow rate

If there was no fault handling, the receiver would sooner or later get empty of liquid, providing bubbling fluid to the evaporators. The evaporator valves would saturate, the heat transfer would drop through the evaporators, and the cabinets would heat up. However, the fault handling strategy is that, if the IT compressors or the receiver valve saturate for a while, the pressure reference is temporarily increased, to get to regions of smaller enthalpy-temperature gradient (which results in more resolute control). Furthermore, this way the convection would increase, as the temperature rises. However, this reduces efficiency significantly, since the high pressure increase through the compressors results in high enthalpy excess.

The problem provided does not have any intuitive ways to go around, therefore it cannot be alleviated by solely a controller specifically designed for both normal operation and fault presence, and at the same time not changing control parameters (for example see the ones presented in [55]). Installing redundant sensor is expensive and does not prevent human error: the two sensors may be installed the same way, and may both experience insulation problems. Therefore, the way to reduce

the negative effects is early fault detection, and proper fault handling, until the fault is fixed.

3.2 State of knowledge on modelling refrigeration systems

The modelling of cooling systems is done today mainly by steady state models, as mostly steady state operation is considered. These are often spatially distributed models, sometimes using finite element method. However, this strategy includes unnecessary high fidelity for a rather dynamically interesting model, but neglects dynamics. The theory on thermodynamical dynamics is immature, including only data driven models (for example in [52]), or models including strong assumptions. However, these latter models often work fine for controller testing. This question was more deeply considered in the special course (see Appendix A), and the results of that will be often restated in this work.

3.3 The dynamical model description

The dynamical model encompasses the high pressure and receiver pressure levels, as described in Section 3.1.1.

3.3.1 Differential equations framework

The derivation of the differential equations was based on the rule for partial derivatives in mathematics, the mass balance equation, the energy balance equation and the low pass filter equation.

When we intend to convert time derivatives of properties to time derivatives of other properties, it is enough to consider the derivatives of two other properties. This can be reasoned by Gibbs' rule [11], or in other words, the phase rule, which states

$$f + N = K + 2, \quad (3.2)$$

where f is the degrees of freedom, N is the number of phases (that is either 1 or 2), and K is the number of components (that is 1). Therefore, we need at maximum two properties to determine a third one. Property conversions refer to converting two properties to a third one, assuming equilibrium. In equilibrium, it will be true for small changes in the properties, that

$$dz = \frac{\partial z}{\partial x}|_y dx + \frac{\partial z}{\partial y}|_x dy. \quad (3.3)$$

The variables x, y, z are properties of thermodynamic systems. By dividing the equation with infinitesimally small time dt , we can determine the necessary time derivatives, when two other ones are provided. Note that the equation is true for any dedicated point in the ph diagram (the partial derivatives are changing with the properties), and assumes equilibrium. This assumption will not hold at any time, since we investigate the dynamical behaviour of thermodynamic systems. However, given that this is the best estimation for the partial derivatives, and that they are changing slowly along the operation domain, it can be safely assume that the results will be reliable.

The mass balance equation of open systems [11] can be stated for a given control volume. It shows that the density in the control volume is increased by inlet mass flow rates and decreased by outlet mass flow rates. The equation is used with zero on the left side, when concentrated elements are considered, since for them no accumulation of mass is assumed.

$$\dot{\rho}V = \sum_j \dot{m}_{in,j} - \sum_k \dot{m}_{out,k} \quad (3.4)$$

The energy balance equation of open systems [11] can be stated for a given control volume. It shows that the energy in the control volume is increased by the inlet enthalpy flow rates (first sum), decreased by the outlet enthalpy flow rates (second sum) and increased by positive heat flow rate \dot{Q} and technical work \dot{W}_t (if it makes work *on* the fluid in the control volume). The equation is used with zero on the left side, when concentrated elements are considered, since for them no accumulation of energy is assumed. Note that potential field (like gravity) and the kinetic energy of the fluid are negligible and not considered.

$$\dot{u}\rho V = \sum_j \dot{m}_{\text{in},j} h_{\text{in},j} - \sum_k \dot{m}_{\text{out},k} h_{\text{out},k} + \dot{Q} + \dot{W}_t \quad (3.5)$$

The low pass filter is the simplest way of creating dynamics for a concentrated element. It is used for realizing actuator dynamics and for damping interactions between states. Given that x is the state and u is the input, the equation states that

$$\dot{x} = \frac{1}{\tau}(-x + u). \quad (3.6)$$

3.3.2 The nonlinear system of differential equations, without simplifications

The system is modelled with a set of differential equations, algebraic constraints and boundary conditions. The boundary conditions are the ambient temperature, the standard air pressure (hence the standard air density ρ_A can be deduced), the enthalpy drop through the heat recovery unit, the enthalpy inlet through from the medium temperature compressor, the heat flow rates from the cabinets, the pressure at the outlet of the receiver valve, and the enthalpy references for the evaporator valves. The evaporator side is modelled as a low pass filter, where the evaporator valves are to provide the correct enthalpy outlet for the evaporators in steady state.

Note that there are some additional boundary conditions introduced for the model. Introducing a new set of differential equations for a smaller control volume (like the volume outlet of the bypass valve) would cause numerical issues due to the fast transients. These are also negligible with regard to the control objectives, hence steady-state values are assumed, and the properties are introduced as inputs after using steady-state property conversions (for example, ρ_{BP}). Therefore these inputs are function of measurements. They can be considered as constraints.

The gas cooler is divided into cells with equal volumes (see Figure 3.3). This is due to the fact that the heat transfer is strongly nonlinear, because of which a spatially continuous or well-discretized model is to be used. One example is that the heat transfer is practically saturated (the maximum is achieved) at the so-called pitch temperature difference, which can occur at any points along the tube. The heat exchange is modelled as counter flow, since although geometrically cross flow is happening, but the same fluid is lead through the serial dimension multiple times, resulting in gradually changing temperature for both media. For the dimensions of the gas cooler, see Figure 3.6 (the picture has been taken from the documentation of TIL Suite library ([click](#))).

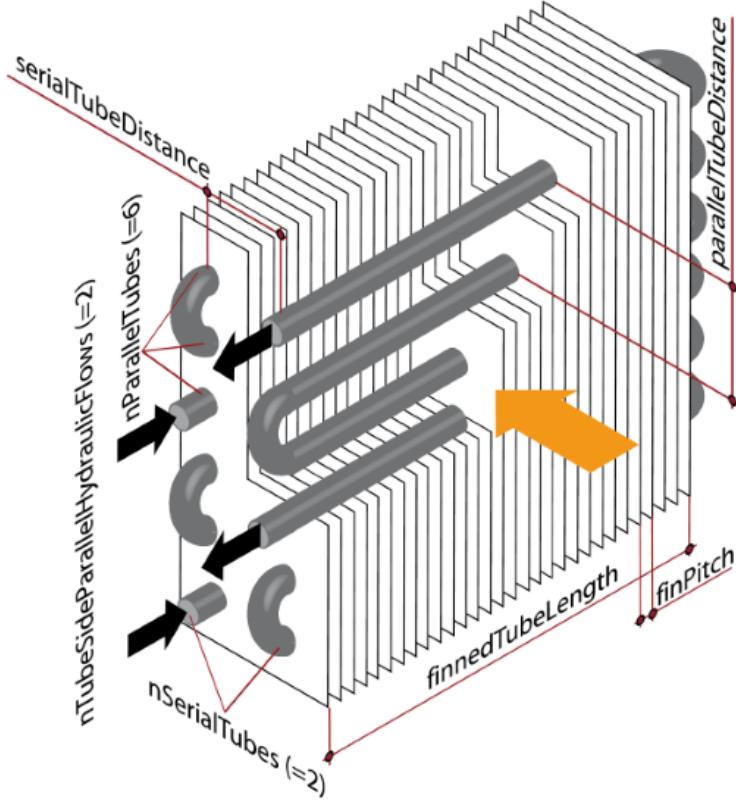


Figure 3.6: Physical structure of a gas cooler (taken from the documentation of TIL Library)

The system can be described in the general form

$$\dot{\mathbf{X}} = \mathbf{f}(\mathbf{X}, \mathbf{U}, \mathbf{p}(\mathbf{X}, \mathbf{U})), \quad (3.7)$$

where x is the state vector, u are the input signals (control outputs) and boundary conditions, $p(x)$ are the parameters being dependent on the states and the inputs. The controller outputs are zero order holded, and the sampling time is 1 second. These can be described as

$$\begin{aligned} \mathbf{X} &= [\dot{V}_A \ p_1 \ h_1 \ \rho_1 \ T_1 \ T_{A,2} \ \dot{m}_{21} \ p_2 \ h_2 \ \rho_2 \ T_2 \ T_{A,1} \ \dots \\ &\quad \text{BP} \ \dot{m}_{\text{HV}} \ p_{\text{R}} \ h_{\text{R}} \ \rho_{\text{R}} \ \dot{m}_{\text{RV}} \ \dot{m}_{\text{IT}} \ \mathbf{X}_{\text{dist}}]^T, \\ \mathbf{X}_{\text{dist}} &= \begin{bmatrix} \dot{m}_{\text{Q}} \\ \Delta h \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} \text{CR}_A \\ \text{CR}_{\text{BP}} \\ \text{CR}_{\text{HV}} \\ \text{CR}_{\text{IT}} \\ \text{CR}_{\text{RV}} \\ \mathbf{U}_{\text{BC}} \end{bmatrix}, \quad \mathbf{U}_{\text{BC}} = \begin{bmatrix} T_{A,0} \\ \rho_{\text{BP}} \\ \rho_G \\ h_G \\ h_L \\ h_{\text{MT}} \\ \Delta h_{\text{HR}} \\ p_{\text{MT}} \end{bmatrix}, \quad (3.8) \\ \mathbf{p}(\mathbf{X}) &= \left[\frac{\partial \rho}{\partial p_1} \quad \frac{\partial \rho}{\partial p_2} \quad \frac{\partial \rho}{\partial p_R} \quad \frac{\partial \rho}{\partial h_1} \quad \frac{\partial \rho}{\partial h_2} \quad \frac{\partial \rho}{\partial h_R} \quad \frac{\partial T}{\partial p_1} \quad \frac{\partial T}{\partial p_2} \quad \frac{\partial T}{\partial h_1} \quad \frac{\partial T}{\partial h_2} \right]^T, \end{aligned}$$

where index A denotes air, index HV high pressure valve, index RV receiver valve, index R receiver, index 0 ambient (air), index MT Medium Temperature, index BP by pass valve opening, index IT parallel compressor, index dist disturbance, index BC boundary condition. Index i denotes the i^{th} cell of the gas cooler with regard to the medium flow direction, having $x_i \in \mathbb{R}^{n_{\text{cell}}}$, $i = 1 \dots n_{\text{cell}}$ (parameter n_{cell} being the number of cells), and $\dot{m}_{i+1,i}$ denotes the mass flow rate from cell i to $i+1$.

Index Q refers to evaporator mass flow rate driven indirectly by heat flow rates from the cabinets, the enthalpy drop Δh_{HR} is the enthalpy drop through the heat recovery unit, and Δh is the enthalpy difference between the virtual enthalpy of the second refrigerant cell and the outlet enthalpy of that cell.

The phenomenon of the virtual enthalpy arises as spatially distributed properties are to be modelled through a given number of volume cells. Enthalpy can be only connected to a certain point along the length, but not to the whole control volume. Infinitely many cell would result in equal inlet and outlet enthalpies, giving a correct model. However, for simplicity, only two cells are used for modelling, leaving the (virtual) enthalpy of a cell badly defined. For the sake of control, this still can be used, as the control objective is connected to a well defined point, the outlet. In order to define the virtual enthalpy (which will not be observable), we may use the weighted average of the cell enthalpies of a high resoluted model, where the weights are the cell densities. This gives a good idea of the average particle enthalpy within a cell.

Two cells are selected for modelling, since experiments from the special course showed that the heat flow rate, being dependent on the temperature difference, has two main phases along the gas cooler tubes. Figure 3.7 presents that at the refrigerant entrance, the temperature difference drops exponentially, then it becomes controlled by the pitch temperature difference (this is a value that so called controls, saturates the maximal heat transfer along the spatial dimension) [11], finally, it is not decaying anymore.

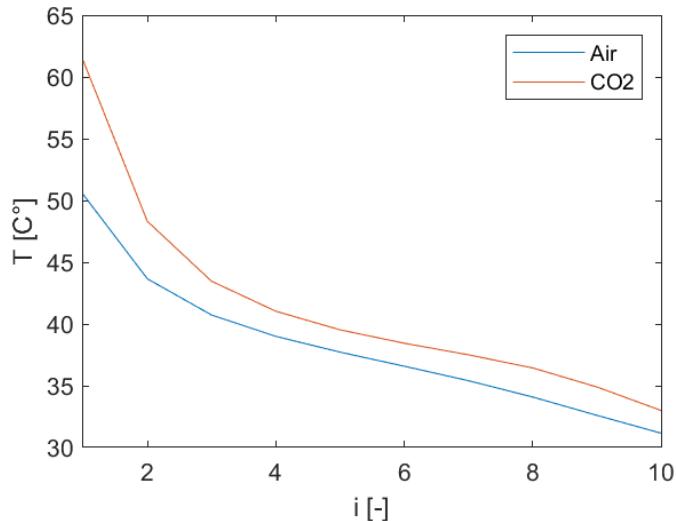


Figure 3.7: Temperature distribution along the cells of the gas cooler, for $n_{cell} = 10$

The fan is controlling the volumetric flow rate of air through the gas cooler with the capacity ratio multiplied by the maximum volumetric flow (this is a significant simplification, since the volumetric flow rate has some hysteresis for low levels of capacity ratio, and saturates above a value):

$$\ddot{V}_A = \frac{1}{\tau_A} \left(-\dot{V}_A + \dot{V}_{A,max} CR_A \right). \quad (3.9)$$

The volumes in the system can be described by the first three equations from Section 3.3.1, with no technical work. The mass balance equation shows that the density is the "least dependent"

property,

$$\dot{\rho}_i = \frac{1}{V_i} \left(\sum_j \dot{m}_{in,i,j} - \sum_k \dot{m}_{out,i,k}, \right), \quad (3.10)$$

while the change of the pressure and the enthalpy are joint. They depend on not only each other (being described by the joint matrix J_i), but on the change in density, and the excitation Ψ as well (see Eq. 3.11). Note that the element $-\frac{p_i}{\rho_i} \dot{\rho}_i$ is a dominating element (resulting in dynamics defining the speed of sound), but is filtered out by the low pass filters in the system, as it describes relatively fast dynamics.

$$J_i \begin{bmatrix} \dot{p}_i \\ \dot{h}_i \end{bmatrix} = \begin{bmatrix} \frac{\dot{\Psi}_i}{V_i} - \frac{p_i}{\rho_i} \dot{\rho}_i \\ \dot{\rho}_i \end{bmatrix} \quad (3.11)$$

$$J_i = \begin{bmatrix} -1 & \rho_i \\ \frac{\partial \rho}{\partial p} \Big|_h & \frac{\partial \rho}{\partial h} \Big|_p \end{bmatrix} \quad \dot{\Psi}_i = \sum_j \dot{m}_{in,i,j} h_{in,j} - \sum_k \dot{m}_{out,i,k} h_{out,k} + \dot{Q}_i$$

The temperature in the cells is directly dependent on any other two properties assuming a property close to equilibrium. The property derivatives already found can be used,

$$\dot{T}_i = \frac{\partial T}{\partial p} \Big|_h \dot{p}_i + \frac{\partial T}{\partial h} \Big|_p \dot{h}_i. \quad (3.12)$$

The change in the mass flow rate between cells is induced by the pressure difference between cells, assuming constant hydraulic resistance ΔR_{hyd} ,

$$\ddot{m}_{i+1,i} = \frac{1}{\tau_{hyd}} \left(-\dot{m}_{i+1,i} + \frac{1}{\Delta R_{hyd}} \sqrt{\rho_i(p_i - p_{i+1})} \right). \quad (3.13)$$

The temperature of the air cell is a weighted average of the inlet air temperature and the refrigerant temperature, where the weight is the ratio of the thermal conductances of the moving air and the convection between the fins of the gas cooler and the air volume,

$$\dot{T}_{A,i} = \frac{1}{\tau_T} \left(-T_{A,i} + \frac{1}{w+1} T_i + \frac{w}{w+1} T_{A,i-1} \right), \quad (3.14)$$

$$w = \frac{\rho_A \dot{V}_A c_p}{\Delta \sigma_{conv}}.$$

The standard air density is denoted by $\rho_{textA} = 1.2 \text{ kg m}^{-3}$, and the isobaric specific heat capacity at the investigated temperature domain is $c_p = 1000 \text{ J (kg K)}^{-1}$. The index conv refers to the convection between the aluminium fins and the air.

The valves are generally described by a maximum opening K_v value, and their mass flow rate is induced by the provided capacity ratio (actuation) and the pressure difference on the two sides (see [34]),

$$\ddot{m}_V = \frac{1}{\tau_V} \left(-\dot{m}_V + K_v \sqrt{\rho_{in}(p_{in} - p_{out})} CR_V \right), \quad (3.15)$$

where index V stands as short for both index HV and index RV.

The IT compressor, described by Eq. 3.16 has a small cabin, in which it compresses the air from the suction to the discharge side. The effective volume of the cabin is called the displacement V_{IT} , and the compressor appears to be a current generator for the system, being constrained by density changed on the suction side. It is controlled by frequency $f_{max} CR_{IT}$. The maximum frequency f_{max}

is 48 Hz. (Note that the necessary torque is not considered in the analysis, but it could be calculated from the technical work in the static version of the energy balance equation, assuming $\dot{Q} = 0$.)

$$\ddot{m}_{IT} = \frac{1}{\tau_{IT}} (-\dot{m}_{IT} + \rho_G V_{IT} f_{\max} \text{CR}_{IT}). \quad (3.16)$$

The bypass valve is also modelled with a low pass filter for the sake of smoothness (the time constant is expected to be as low as for the other valves),

$$\ddot{m}_{BP} = \frac{1}{\tau_{BP}} (-\dot{m}_{BP} + \text{CR}_{BP} \dot{m}_{HV}). \quad (3.17)$$

The disturbances have zero derivatives, since they are supposed to change slowly, and there is no exact way to dynamically investigate them, hence

$$\begin{aligned} \dot{\Delta h}_2 &= 0, \\ \dot{m}_Q &= 0. \end{aligned} \quad (3.18)$$

The derivations of the equations of this section can be found in Section 3.3.3.

3.3.3 Derivation of nonlinear differential equations, and constraints

In order to find an expression of the time derivative of the energy in terms of the properties, we use the definition of enthalpy $h = u + pv$ and the chain rule when differentiating the specific volume $v = \frac{1}{\rho}$. Letter d denotes the differential operator.

$$\begin{aligned} \rho V \dot{u} &= \dot{\Psi} \\ \dot{u} &= \frac{d(h - pv)}{dt} = \dot{h} - \frac{1}{\rho} \dot{p} + \frac{p}{\rho^2} \dot{\rho} \end{aligned} \quad (3.19)$$

The partial derivatives in Equation 3.11 are derived using the chain rule of partial differentiation,

$$\dot{\rho} = \frac{\partial \rho}{\partial p} \Big|_h \dot{p} + \frac{\partial \rho}{\partial h} \Big|_p \dot{h}. \quad (3.20)$$

Similarly to the previous states, the temperature states of the gas cooler can be assigned directly to any other two ones, assuming equilibrium, giving Eq 3.12.

To deduce the interaction between the refrigerant and the air cells, four stages of thermal resistance shall be considered. Approaching from the refrigerant to the gas inlet of a cell, these are the following.

- The convection between the refrigerant and the copper tube wall.
- The resistance of the copper tube and the aluminium fins.
- The convection between the aluminium fins and the gas (air).
- The heat resistance of air, described by the (static) energy balance equation of air.

In practice, the first two are negligible compared to the last two, there are more orders of magnitude differences in thermal resistance. This is why

$$\frac{1}{\sigma} = \sum_{j=1}^4 \frac{1}{\sigma_j} \approx \frac{1}{\sigma_{\text{conv}}} + \frac{1}{\rho_A \dot{V}_A c_p}. \quad (3.21)$$

The heat transfer to a cell \dot{Q}_i and the convection σ_{conv} between the fins and the air volume are calculated by the equations below. Note that the convection is assumed to be the same all along

the tube, and that the convection is assumed to be the linear function of volumetric flow rate, with coefficients σ_0 (natural convection) and k_{conv} (forced convection slope). The heat transfer can be imagined as two serially connected resistances, which both depend on the volumetric flow rate as seen in Figure 3.8.

$$\begin{aligned}\dot{Q}_i &= (T_{A,i} - T_i) \Delta\sigma_{\text{conv}} \\ \Delta\sigma_{\text{conv}} &= \frac{\sigma_{\text{conv}}}{n_{\text{cell}}}, \\ \sigma_{\text{conv}} &= \sigma_{\text{conv},0} + k_{\text{conv}} \dot{V}_A.\end{aligned}\quad (3.22)$$

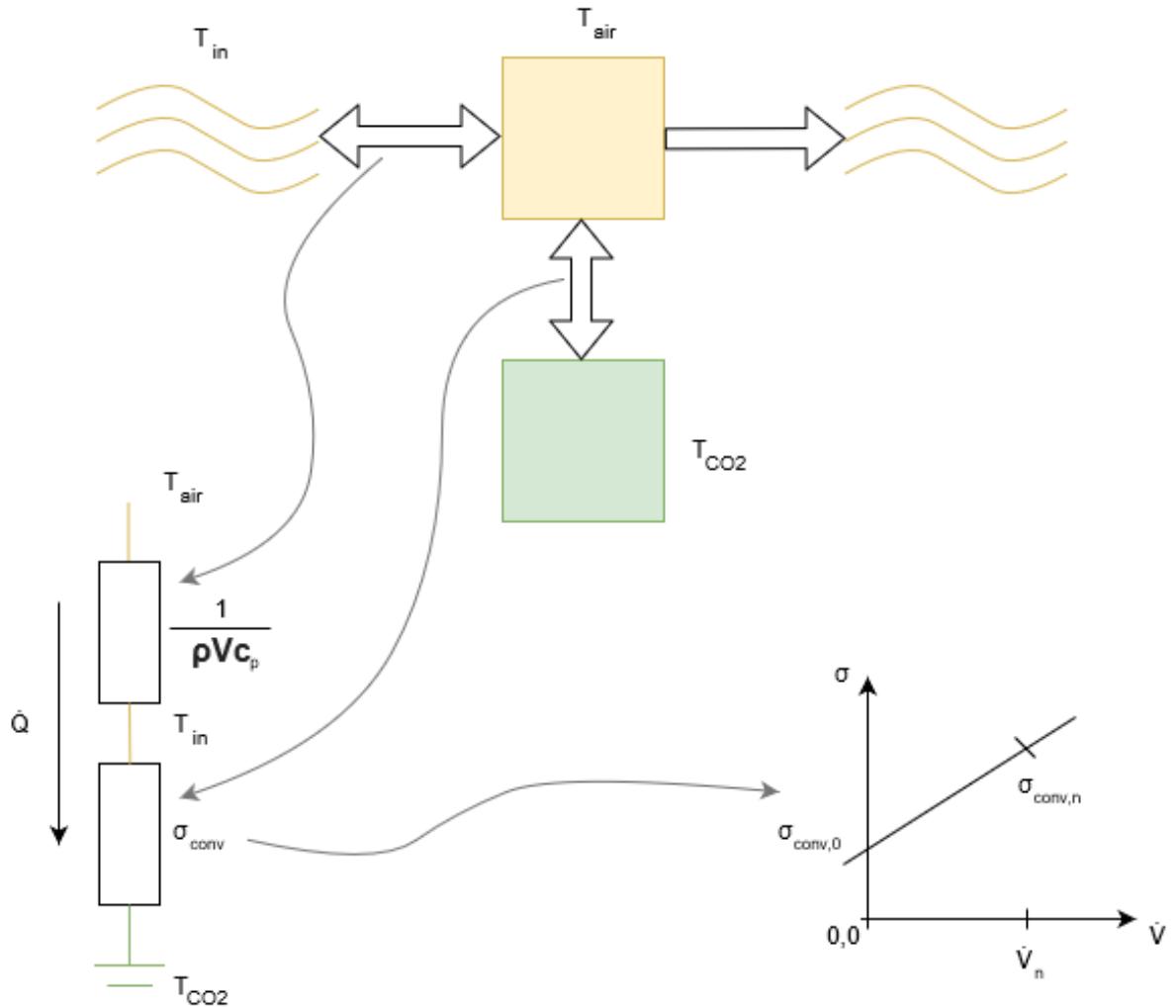


Figure 3.8: Modelling of heat transfer. Physical interactions (up), electrical analogy (left), convection model (right)

For the two convection parameters, the two equations in Eq. 3.23 are used. Index nom denotes nominal quantities. The first one uses the arbitrary design parameter $\alpha = 5$ (which is taken as a parameter used to produce reasonable simulations in the Modelica model). The second equation approximates the nominal convection from the nominal heat flow rate $\dot{Q}_{\text{nom}} = 74.3\text{kW}$, and the pitch temperature difference. Not having more information about the distribution, the latter is assumed to be at the refrigerant outlet, being $\Delta T_{\text{pitch}} = 3\text{ K}$. Note that in case of exponential curves, $\Delta T_{\log} \equiv \frac{\Delta T_{\text{out}} - \Delta T_{\text{in}}}{\log \frac{\Delta T_{\text{out}}}{\Delta T_{\text{in}}}}$ would be another way of approximating the temperature difference. However, the considered curves are not exactly exponential, and according to simulation tests on

the high fidelity model described in Section 5.1, the value $\Delta T_{log} \approx 20 K$ gives around 10% lower heat transfer than the required one.

$$\begin{aligned}\sigma_{conv,0} &\approx \frac{\sigma_{conv,n}}{\alpha} \\ \sigma_{conv,n} &= \frac{\dot{Q}_{nom}}{\Delta T_{pitch}} = \sigma_{conv,0} + k_{conv} \dot{V}_{A,nom}\end{aligned}\quad (3.23)$$

For spatial discretization, the manner of the heat conduction was considered. Assume that the convection is the product of the area A_{fin} and the heat transfer coefficient U , the latter being only dependent on the materials of fin and air, then having equal heat transfer areas for every cell gives

$$\begin{aligned}\sigma_{conv} &= U A_{fin}, \quad \text{hence} \\ \Delta\sigma_{conv} &= U \Delta A_{fin} = \frac{\sigma_{conv}}{n_{cell}}.\end{aligned}\quad (3.24)$$

Two further equations are used to derive the results of Equation 3.22, the (static) energy balance equation for air (derived from the definition of the specific heat capacity $c_p \equiv \frac{Q}{\Delta T_m}$),

$$\dot{Q}_i = (T_{A,i-1} - T_A) \rho_A \dot{V}_A c_p + \frac{d(T_{A,i-1} - T_A)}{dt} \rho_A \Delta V c_p, \quad (3.25)$$

and the assumption, that the heat flow rate can be approximated as in electronics, as the product of the temperature difference (intensive variable or potential difference) and the convection (conduction), and assuming some capacitance,

$$\begin{aligned}\dot{Q}_i &= (T_{A,i} - T_i) \sigma_{conv} + \frac{d(T_{A,i-1} - T_A)}{dt} \tau_A, \\ \tau_A &\equiv \frac{\Delta V_A}{\dot{V}} = \frac{A_{GC} H_{GC}}{n_{cell} \dot{V}_A},\end{aligned}\quad (3.26)$$

where A_{GC} is the horizontal area and H_{GC} is the height of the gas cooler, n_{cell} is the number of cells chosen to model the gas cooler. For simplicity, and in the lack of literature about heat flow rate dynamics, the time constant in Equation 3.26 is assumed to be the same as the time constant in Equation 3.25. Note that this is a very small value (≈ 1 second), and will be modified to 150 seconds based on cross-covariance investigation in Section 7.4.

The receiver is insulated, so the heat transfer to the surrounding is assumed to be zero, $\dot{Q} = 0$ in the energy balance equation. There are three mass flow rates, the inlet, and the gas and the liquid outlets. The outlet enthalpies and densities are read from the ph diagram, assuming equilibrium, which is a well-based assumption as the receiver can be considered as a buffer.

For the valves, it can be deduced from the static version of the energy balance equation (assuming no technical work and heat flow rate), that the outlet enthalpy is the same as the inlet enthalpy. Note that the outlets are not explicitly denoted.

$$h_V = h_{V,in} \quad (3.27)$$

An ideal compressor raises pressure by staying at the same entropy, then the ph diagram could be used to find the outlet enthalpy [11]. However, the isentropic efficiency $\eta_s \approx 0.6$ is to be considered as a value from the industry (for a particular *Bitzer* compressor) [15].

$$h_{comp} = h_{comp,in} + \frac{h_{comp,id}(h_{comp,in}, s_{in}) - h_{comp,in}}{\eta_s} \quad (3.28)$$

The fluid dynamics in the gas cooler (used in Eq. 3.13) are considered how it is described in the chapter *Dimensioning of tubes* in [18]. Even with rough estimations, it turns out that the flow is

turbulent. The estimation of the friction factor f of the pipe is done by iterating the definition for Reynold's number (denoted as Re), the Gnielinski correlation, and the Darcy–Weisbach equation, until convergence is achieved. Note that the range of Gnielinski correlation is $2300 < \text{Re} < 5e6$. The iterated equations are

$$\begin{aligned}\text{Re} &= \frac{4\dot{m}}{\pi d\mu}, \\ f &= \frac{0.5}{(0.79 \log \text{Re} - 1.64)^2}, \\ R_{\text{hyd}} &= \sqrt{\frac{16fl}{d^5\pi^2}}, \\ \dot{m} &= \frac{1}{R_{\text{hyd}}} \sqrt{\rho\Delta p},\end{aligned}\tag{3.29}$$

where d is the diameter, l is the length, Δp is the pressure difference along the two end points of the tube, the dynamic viscosity is given by $\mu = \rho\nu$, and the kinematic viscosity ν is the property of the fluid. Then the last equation is used to determine the mass flow rate between cells. The discretization is done by considering that only the length of a tube slice changes by altering the number of cells,

$$\Delta R_{\text{hyd}} = \sqrt{\frac{16f\Delta l}{d^5\pi^2}} = \sqrt{\frac{16fl}{d^5\pi^2 n_{\text{cell}}}}.\tag{3.30}$$

The outlets of the IT and the MT compressors are joined, then inputted to the heat recovery unit, which then passes it to the gas cooler. The following joint equations can be deduced from the balance equations (Eq. 3.4 and Eq. 3.5), assuming no accumulation of mass or energy. In the following equation, every index refers to the outlet of a component.

$$\begin{aligned}\dot{m}_{\text{HR}} &= \dot{m}_{\text{MT}} + \dot{m}_{\text{IT}}, \\ h_{\text{HR}} &= \frac{h_{\text{MT}}\dot{m}_{\text{MT}} + h_{\text{IT}}\dot{m}_{\text{IT}}}{\dot{m}_{\text{HR}}} - \Delta h_{\text{HR}}.\end{aligned}\tag{3.31}$$

The bypass valve is modelled as a double current generator (see Figure 3.9), which linearly distributes the flow according to a control signal (note that in reality valve equations are quadratic for mass flow rate). On one side only a resistance, on the other, a resistance and a capacitance are considered. The time constant of the condensator is τ_{hyd} , a custom value. The inlet enthalpy of the two flows are the same. However, the outlet enthalpy is defined by the balance equations again. The pressure at the outlet is assumed to be the same, as the pressure at the last refrigerant cell of the gas cooler.

$$\begin{aligned}\dot{m}_{\text{BP}} &= \dot{m}_{\text{HV}}\text{BP} \\ \dot{m}_1 &= \dot{m}_{\text{HR}} - \dot{m}_{\text{BP}} \\ \dot{m}_2 &= \dot{m}_{\text{HV}}(1 - \text{BP}) \\ h_{\text{BP}} &= \frac{\dot{m}_2(h_2 - \Delta h) + \dot{m}_{\text{BP}}h_{\text{HR}}}{\dot{m}_{\text{HV}}}\end{aligned}\tag{3.32}$$

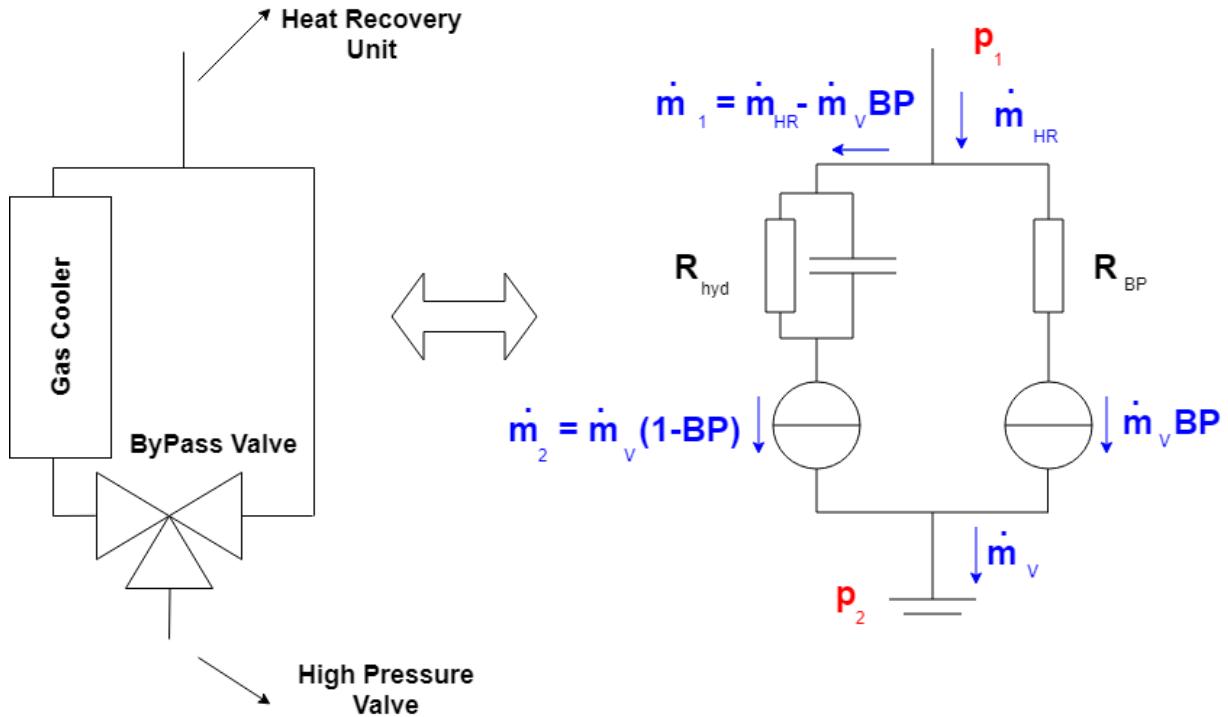


Figure 3.9: Modelling of the bypass valve. Physical model (left), electrical analogy (right)

3.3.4 Measurements and constraints

The measurement equations are defined as

$$\mathbf{Y} = \mathbf{g}(\mathbf{X}, \mathbf{U}, \mathbf{p}(\mathbf{X})), \quad (3.33)$$

where \mathbf{Y} are the measurements. It can be extracted to

$$\mathbf{Y} = \begin{bmatrix} p_{2m} \\ h_{BPm}(p_{2m}, T_{BPm})_{ss} \\ p_{Rm} \\ h_{Rm}(p_{Rm}, x_{Rm})_{ss} \\ h_{HRm}(p_{1m}, T_{HRm})_{ss} \end{bmatrix} = \begin{bmatrix} p_2 + v_{p2} \\ h_{BPm}(p_2 + v_{p2}, T_{BP} + v_{TBP})_{ss} \\ p_R + v_{pR} \\ h_{Rm}(p_R + v_{pR}, x_R + v_{xR})_{ss} \\ h_{HRm}(p_1 + v_{p1}, T_{HR} + v_{THR})_{ss} \end{bmatrix}, \quad (3.34)$$

where the operation $c(a, b)_{ss}$ refers to the property conversion from a and b to c , assuming equilibrium. The index m refers to being a measurement in the model, adding proper variance Gaussian white noise to the original physical measurements. They are defined as

$$v_i \in \mathcal{N}(\mu_i, \sigma_i^2), \quad (3.35)$$

where v_i is the superimposed stochastic variable with mean μ_i standard deviation σ_i , and variance σ_i^2 . The variances are taken as the maximum variances of inspected industrial measurements. Since there are only pressure and temperature sensors, and a liquid level sensor in the receiver, they are described by property conversions. Note that these property conversions assume equilibrium, and as the conversions are nonlinear, the output properties are not normal distributed. The measurement x_{Rm} is the liquid level (measured in meters) in the receiver, and the way to convert it to vapour quality (vapor ratio in the two-phase fluid) is described by the following equation; then

property conversion can be applied.

$$q = \frac{\rho_{\text{vapour}}(1-x)}{\rho_{\text{vapour}}(1-x) + \rho_{\text{liquid}}x} \quad (3.36)$$

Lacking a measurement of \dot{m}_Q would go against observability. The boundary conditions assume constant heat flow rate into the system (people are opening the door of the cooler and the freezer cabinets constantly, and the changing temperature difference is neglected here). Index C denotes cooler outlet, index F freezer outlet. These values are the reference enthalpies for control, as perfect control is assumed. This is a reasonable assumption, since the evaporator dynamics are swifter, than the ones in the back-end.

$$\begin{aligned}\dot{m}_F &= \frac{\dot{Q}_F}{h_F - h_L} \\ \dot{m}_C &= \frac{\dot{Q}_C}{h_C - h_L} \\ \dot{m}_Q &= \dot{m}_C + \dot{m}_F\end{aligned}\quad (3.37)$$

Note that in this set-up, the state combination of Δh and the enthalpies in the gas cooler is not observable, and leaves integrators in the non-observable subspace. This issue is going to be addressed in Section 3.4.

3.4 Simplifications of the model

The nonlinear model derived in Section 3.3.2, although very detailed, is very impractical to use for estimation or control. First the reasons for the simplifications, then the simplifications themselves will be described, finally the new model will be stated as a whole.

3.4.1 Reasons for simplifications

The reasons for simplifications were the following:

- There were negligible details in the model. It included different time scales, namely thermodynamics, actuator dynamics and in the very original model, sound dynamics. The thermodynamics involve time constants above ≈ 10 seconds, while the actuator dynamics or the tube resistance range up to maximum some seconds. The other example is that, in Eq. 3.11, different magnitudes are involved. According to experiments from the special course, the term including \dot{p} is two order of magnitude smaller than the term including \dot{h} .
- Large parameter uncertainty was also an issue. Actuator dynamics or flow resistance are not provided by data sheets, and estimations would require additional measurements for applying in the industry.
- During the linear analysis, which will be carried out for the simplified model in Section 4, it was found that there is a larger non-controllable subsystem with integrators. This is due to the joint relation of three derivatives in Eq. 3.11, presented as two equations, which constrains the model. It can be visualized as pouring water out of a glass into another one: both rates of change of the volumes can be controlled, however, they are constrained by the flow rate. The motivation to prevent this problem is to ease the job of a potential model based controller, by providing well-scaled controllability Grammians [21].
- There is lack of observability as well. This is due to the fact, that the states within the gas cooler cannot be measured (partly because of feasibility, partly due to their virtual nature).

- In case of too many states and too few measurements, the estimator starts to distribute the information from the innovation signal incorrectly. The more indirectly observed states, the more difficult to tune the estimator. This is worsened by the fact that the simulated Gaussian noise is added to the temperature measurement, which is then converted to enthalpy 'measurement'. The resulting signal is not guaranteed to be Gaussian anymore.
- There were numerical issues connected to the model. Partial differentials change significantly along the operation domain, sometimes several order of magnitudes (see Figure 3.10). If a poorly estimated state is multiplied by a non-fitting value from the look-up table of the partial differentials, the estimation errors grow in time.

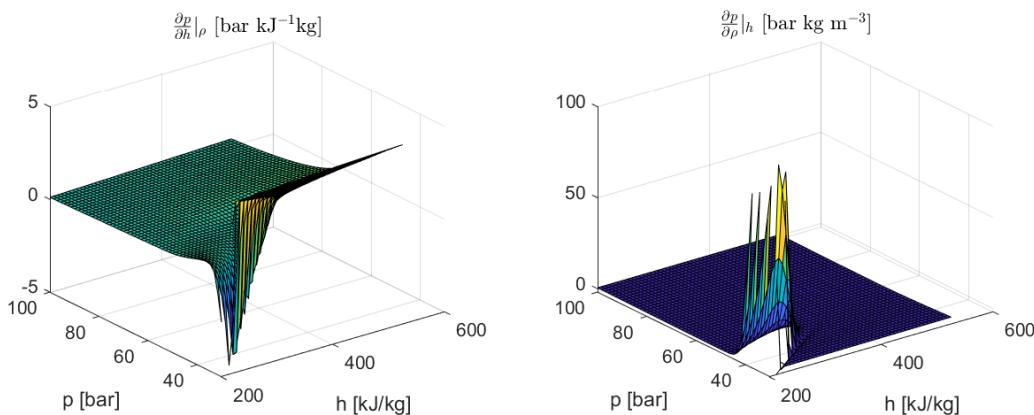


Figure 3.10: Partial differentials along a pressure-enthalpy grid

3.4.2 Simplifications

The ways to overcome the problems in Section 3.4.1 were the following:

- Instead of connecting the derivatives with partial differentials, they were connected by well-resolved least squares estimations along the operation domain (see Section 4.2). These provide smoother surfaces, since they are supposed to be the integrands of the partial differentials. Note that this is a further simplification: until now only the partial differentials assumed equilibrium, however, now the property values are also defined in equilibrium (which condition is not fulfilled, as a dynamical model is investigated).
- In order to go around the interdependence of the properties, and that two of them are eligible to determine the rest in equilibrium, the perception of the system has been changed. Since the balance equations reflect on the change of the density and the enthalpy of the system, they are considered from now on as 'true properties', and pressure and temperature become the 'projected properties' of them. In order to implement this view, the \dot{p} term is neglected from Eq. 3.11, disjointing the two equations. The pressure state is then connected to its least squares estimation from pressure and enthalpy, through a low pass filter with a pressure dynamic time constant, which, neglecting sound dynamics, can be maximized as ≈ 0.1 second. This way, the controllability problem is solved.
- Only one cell was considered in the gas cooler. This abolished the usability of the features of the spatial temperature difference distribution, but made the model less error prone to the parameter uncertainty introduced by the tube resistance model. Significantly less states also eased the problem of information distribution.
- The issue with observability is solved, and the issue with information distribution is eased by providing constraints as measurements. These constraints use physical quantities at the

refrigerant and gas inlets of the gas cooler as input, to find the virtual temperature or enthalpy. The method used was linear least squares. To be less dependent on model parameters, dynamics are not considered. This sets the accuracy of the estimations to the accuracy of the constraints. Delays were estimated, as described in Section 7.4. Nevertheless, the results were very sensitive to model parameters, therefore they did not work on field data. Finally, based on experiments (see Figure 3.11), it was found that the weighted average of the enthalpies of the cells can be approximated by $h_{GC} \approx \frac{1}{3}h_{HR} + \frac{2}{3}h_{BP}$, where h_{GC} is the virtual enthalpy of the gas cooler.

- No actuator dynamics are considered, as they were negligible, and it eased the problem of information distribution of the estimator.
- The field data were recorded without bypass valve and heat recovery unit, at around 20 degrees. Therefore the scope of the project was reduced from 10 to 30 °C, where these items are not used.
- Since there is no way to directly estimate \dot{m}_Q , in order to ease the problem of information distribution, it is introduced as an input, instead of being a measurement.

After applying these simplifications, the (linearized and discrete time, see later chapters) model based estimations are compared to the results of a high fidelity model in Section 9.3. Controllability analysis is done to the linearized model in Section 4.1.3 as well. These model evaluations motivated no further simplifications or change of the model.

3.4.3 Simplified nonlinear differential equations

After the simplifications, Eq. 3.7, the states, the inputs and the parameters become

$$\mathbf{f}(\mathbf{X}) =$$

$$\left[\begin{array}{c} \frac{\rho_{GC} \partial_{pp} GC - p_{GC} + \partial_{ph} GC h_{GC}}{\tau_p} \\ \frac{h_{HR} (\dot{m}_Q + CR_{IT} f_{IT,max} V_{IT} \rho_G) + CR_{HV} K_{v,HV} \sqrt{\rho_{BP} (p_{GC} - p_R)} (\delta_h - h_{GC}) + CR_A \dot{V}_{A,max} c_p \rho_A (T_{A,0} - T_{A,1})}{V_{cell} \rho_{GC}} \\ \frac{\dot{m}_Q - CR_{HV} K_{v,HV} \sqrt{\rho_{BP} (p_{GC} - p_R)} + CR_{IT} f_{IT,max} V_{IT} \rho_G}{V_{cell}} \\ \frac{\rho_{GC} \partial_{T\rho} GC - T_{A,1} + T_{A,0} w - T_{A,1} w + \partial_{Th} GC h_{GC}}{\tau_A (w+1)} \\ \frac{\rho_R \partial_{pR} - p_R + \partial_{ph} R h_R}{\tau_p} \\ \frac{\dot{m}_Q h_L + CR_{RV} K_{v,G} h_G \sqrt{-\rho_G (p_{MT} - p_R)} + CR_{HV} K_{v,HV} \sqrt{\rho_{BP} (p_{GC} - p_R)} (\delta_h - h_{GC}) + CR_{IT} f_{IT,max} V_{IT} \rho_G h_G}{V_R \rho_R} \\ - \frac{\dot{m}_Q - CR_{HV} K_{v,HV} \sqrt{\rho_{BP} (p_{GC} - p_R)} + CR_{RV} K_{v,G} \sqrt{-\rho_G (p_{MT} - p_R)} + CR_{IT} f_{IT,max} V_{IT} \rho_G}{V_R} \\ 0 \end{array} \right]$$

$$\mathbf{X} = \begin{bmatrix} p_{GC} \\ h_{GC} \\ \rho_{GC} \\ T_{A,GC} \\ p_R \\ h_R \\ \rho_R \\ \mathbf{X}_{dist} \end{bmatrix}, \quad \mathbf{X}_{dist} = \Delta h_{GC}, \quad \mathbf{U} = \begin{bmatrix} CR_A \\ CR_{HV} \\ CR_{IT} \\ CR_{RV} \\ \mathbf{U}_{BC} \end{bmatrix}, \quad \mathbf{U}_{BC} = \begin{bmatrix} \dot{m}_Q \\ \rho_{BP} \\ \rho_G \\ h_G \\ h_L \\ h_{HR} \\ p_{MT} \\ T_{A,0} \end{bmatrix},$$

$$\mathbf{p}(\mathbf{X}) = \begin{bmatrix} \partial_{ph,GC} \\ \partial_{ph,R} \\ \partial_{p\rho,GC} \\ \partial_{p\rho,R} \\ \partial_{Th,GC} \\ \partial_{T\rho,GC} \end{bmatrix}, \quad (3.38)$$

where index GC refers to the virtual states of the gas cooler, and the ∂ symbol refers to parameters in the least squares estimation $c = \partial_{ca} a + \partial_{cb} b$. Note that these parameters depend on the actual operation point, just like the partial differentials used to do.

Due to the simplifications, Eq. 3.11 has become

$$\dot{h} = \frac{1}{V_i \rho_i} \Psi_i \quad (3.39)$$

and

$$\dot{p}_i = \frac{1}{\tau_p} (-p_i + \partial_{ph,i} h_i + \partial_{p\rho,i} \rho_i). \quad (3.40)$$

The refrigerant temperature state has been dropped, since due to the simplifications, the derivative of it is not relevant anymore, and least squares is eligible to calculate it,

$$T_i = \partial_{Th,i} h_i + \partial_{T\rho,i} \rho_i. \quad (3.41)$$

3.4.4 Measurements and constraints

After the simplifications, the measurements in Eq. 3.33 become

$$Y = \begin{bmatrix} p_{GCm} \\ h_{BP,m}(p_{GCm}, T_{BP,m})_{ss} \\ p_{Rm} \\ h_{Rm}(p_{Rm}, x_{Rm})_{ss} \\ h_{GCm} \end{bmatrix}, \quad (3.42)$$

where the measurement of the virtual enthalpy is found by the rule of thumb

$$h_{GCm} = \frac{1}{3}h_{HR} + \frac{2}{3}h_{BP}, \quad (3.43)$$

which is inspired by simulations, where for $n_{cell} = 10$, the virtual enthalpy h_{GC} is approximated by the following equation, and is depicted in Figure 3.11. The value of h_{HR} is delayed according to Section 7.4. The numbers in Eq. 3.43 may seem wrong after looking at the figure, however, due to the steep initial temperature drop of the refrigerant, the true inlet enthalpy is hidden, which is around 500 kJkg^{-1} . Note that in this diagram, between 5000 seconds and 14000 seconds, the fault is introduced, causing gas loop. Hence the enthalpy difference between the cells drop, indicating less effective heat transfer, and resulting in larger outlet enthalpy.

$$h_{GC} = \frac{\rho_i h_i}{\sum_i \rho_i}. \quad (3.44)$$

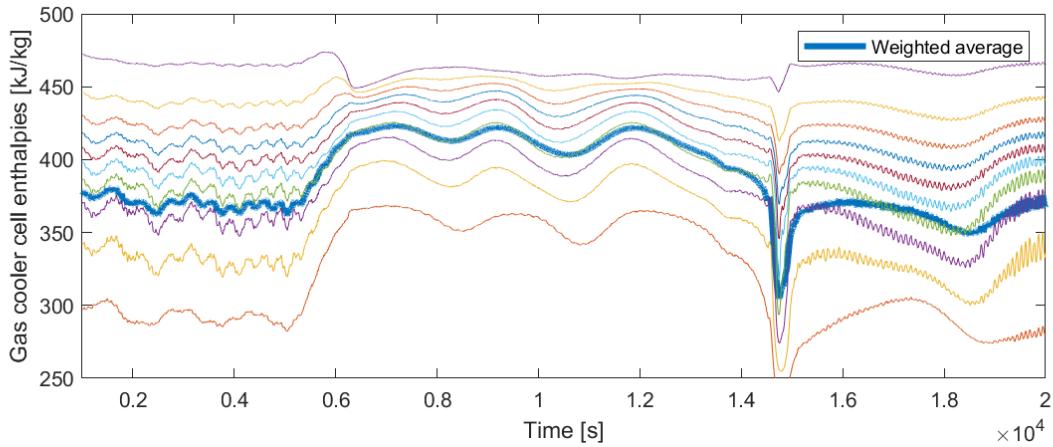


Figure 3.11: Gas cooler cell enthalpies for equal volume cells for a 10-cell simulation

The expressions of Eq. 3.37 are rough estimates, and were not providing sufficient values for the state estimator. At that stage, it was assumed that there would be no measurements at the middle-temperature pressure level. However, to solve this problem the pressure and the temperature (converted to density) at the inlet of the MT compressors have become incorporated to the model. Therefore the disturbance input \dot{m}_Q is expressed as

$$\dot{m}_Q = \rho_{MT,in} V_{MT} f_{max} CR_{MT}, \quad (3.45)$$

which provides better estimation for the high pressure, and worse for the receiver pressure estimates.

3.4.5 The look-up tables for property conversions

The transfer parameters, denoted by ∂ , were addressed to their values along a grid in the operation domain, with linear least squares. This meant, that around every resolution point, a further resolution was created, and the estimations of pressure and temperature were given for every point

$$\hat{Y} = \partial_{\hat{Y}h} h + \partial_{\hat{Y}\rho} \rho, \quad (3.46)$$

where \hat{Y} is the estimation of the output, which is either temperature or pressure. The sum of squares of the estimation errors is minimized by the parameter choice of ∂ , if the following formula is used ([5]),

$$\begin{bmatrix} \partial_{\hat{Y}h} h \\ \partial_{\hat{Y}\rho} \rho \end{bmatrix} = (\boldsymbol{\theta}^T \boldsymbol{\theta})^\dagger \boldsymbol{\theta} \hat{\mathbf{y}}, \quad (3.47)$$

where $\boldsymbol{\theta}$ is the vector of enthalpy-density pairs, and $\hat{\mathbf{y}}$ is the vector of outputs. The resolution was chosen as 200 equal intervals, then least squares of 5 near points to the target points. The operation domain is 30...100 bar, and 200...525 kJkg⁻¹. The map of the transfer parameters can be seen in Figure 3.12. The largest errors along are grid is 0.66 °C, and 0.2 bar, which are eligible for the target application.

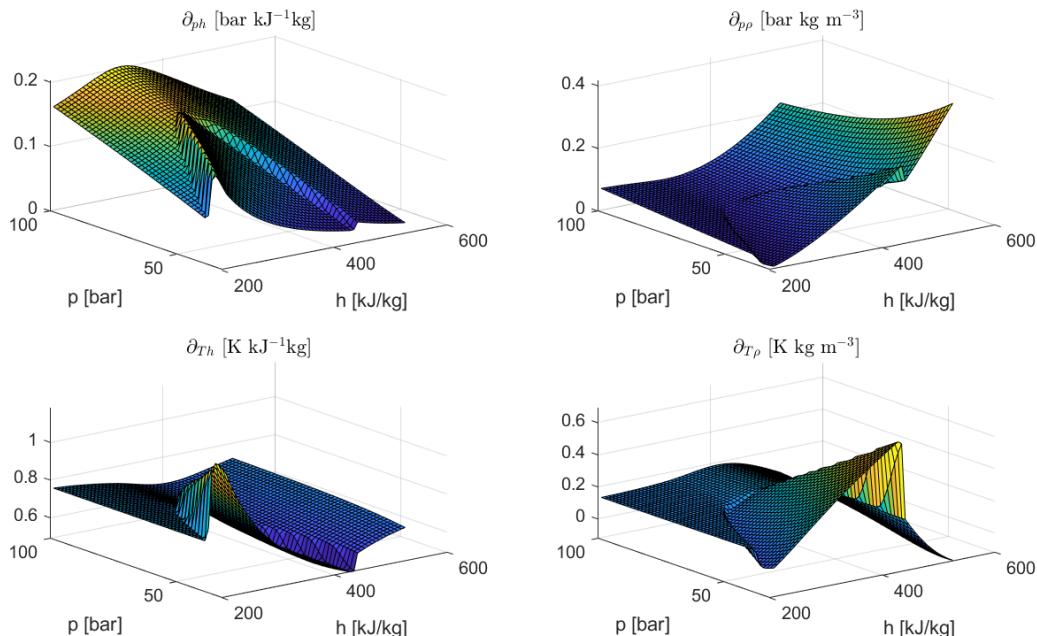


Figure 3.12: Linear least squares transfer parameters along a pressure-enthalpy grid

Chapter 4

Linear System Analysis

Chapter 3 provides a continuous time nonlinear system, which is linearized below. This serves two goals: the first is to have a computationally easy and predictable problem for estimation - since in case of implementation, the code is to be run on a microcontroller, PLC, or other field unit. The second is to have to have a proper form for stability, controllability, observability analysis and controller design, as the literature behind linear analysis is mature and involves closed form solutions. The linearization parameters depend on the set-point, therefore a linear time varying (LTV) estimator is to be used. However, these dynamics do not change rapidly along the operation domain, therefore the linearization itself shall be safely applied. In this chapter, first a well-defined linearization point will be investigated, which is, due to its supercritical nature, is important in refrigeration system design. This is where the coefficient of performance (COP, see 9.2.1) is the lowest. Afterwards, the issues of analysing the linear time varying system is described.

4.1 Linear time-invariant system at a well-defined linearization point

The chosen linearization point involves a receiver state under the saturation curve (see Figure 3.1), while gas cooler states over the saturation curve. This is the biggest challenge for carbon-dioxide refrigeration system, since the required heat transfer is difficult to achieve. For this reason, several data sheets for gas coolers, or refrigeration back-end systems, include the description of this operation point, to show the occurring COP value (which will be described in Section 9.2.1). The linearization is done by Taylor-expanding Eq. 3.7 and Eq. 3.33 around the linearization point and neglecting the square or the higher order terms, assuming that the parameters change slowly,

$$\begin{aligned} \mathbf{f}(\mathbf{X}, \mathbf{U}, \mathbf{p}(\mathbf{X})) &\approx \mathbf{f}(\mathbf{X}_{ss}, \mathbf{U}_{ss}) + \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \\ \mathbf{x} &= \mathbf{X} - \mathbf{X}_{ss}, & \mathbf{u} &= \mathbf{U} - \mathbf{U}_{ss}, \\ \mathbf{A} &= \frac{\partial f}{\partial \mathbf{X}}(\mathbf{X}_{ss}, \mathbf{U}_{ss}) & \mathbf{B} &= \frac{\partial f}{\partial \mathbf{U}}(\mathbf{X}_{ss}, \mathbf{U}_{ss}), \\ \mathbf{g}(\mathbf{X}, \mathbf{U}, \mathbf{p}(\mathbf{X})) &\approx \mathbf{g}(\mathbf{X}_{ss}, \mathbf{U}_{ss}) + \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}, \\ \mathbf{y} &= \mathbf{Y} - \mathbf{Y}_{ss}, \\ \mathbf{C} &= \frac{\partial g}{\partial \mathbf{X}}(\mathbf{X}_{ss}, \mathbf{U}_{ss}) & \mathbf{D} &= \frac{\partial g}{\partial \mathbf{U}}(\mathbf{X}_{ss}, \mathbf{U}_{ss}), \end{aligned} \tag{4.1}$$

where \mathbf{A} is the system matrix, \mathbf{B} is the input matrix, \mathbf{C} is the measurement matrix, \mathbf{D} is the feedforward matrix, \mathbf{X}_{ss} is the state at the linearization point, \mathbf{U}_{ss} is the input at the linearization point. From now on, the dynamical equations of

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}, \end{aligned} \tag{4.2}$$

are called the linear continuous time state space model. From the end of this chapter, they are denoted by index cont, since after being discretized in Section 6.1, they are not used anymore.

However, the analysis can be done on the continuous system, since sampling time is not a challenge at the estimation and control of refrigeration systems.

The nonlinear differential equations in Eq. 3.38 were implemented in Matlab as symbolic expressions. Then, in order to linearize them, the JACOBIAN function was used.

There were four types of analyses carried out on the linear time invariant system. Beforehand, in order to be able to interpret the numbers, the states were divided by the largest tolerable deviations from their control references; this can be perceived as normalization and is solved by state transformations. The analyses were the following, and were visualised by colour coding. For visualization reasons, the matrices expressing dynamics are depicted after taking the piecewise exponentials of their values.

- Investigation of the state interactions through the system matrix.
- Eigenvalue analysis and the state contributions to the eigenvalues.
- Controllability staircase form [21].
- Observability staircase form.

For the last three analysis, transformations are needed. When $\mathbf{x}' = \mathbf{T}^{-1}\mathbf{x}$ is defined by the transformation matrix \mathbf{T} , the new state space model is equivalent to the previous one dynamically, just describes the other states with the following state space matrices:

$$\begin{aligned} \mathbf{A}' &= \mathbf{T}\mathbf{A}\mathbf{T}^{-1}, \\ \mathbf{B}' &= \mathbf{T}\mathbf{B}, \\ \mathbf{C}' &= \mathbf{C}\mathbf{T}^{-1}, \\ \mathbf{D}' &= \mathbf{D}. \end{aligned} \tag{4.3}$$

4.1.1 The system matrix

In Figure 4.1, the colour coded piecewise exponential of the normalized system matrix can be seen. Note that the illustrated system matrix is transformed, letting every state value be normalized by its maximum deviation value (assuming that the controller is working well). Piecewise exponentials are used to find an adequate domain for visualization. Here, some important interactions can be observed, like

- the pressure decreases itself (since it increases mass flow rate through the valve),
- the gas cooler temperature decreases itself (due to the increasing heat flow rate),
- enthalpy levels increases pressure levels,
- the disturbance is an integrator,
- the pressure and the enthalpy in the gas cooler directly increase the enthalpy and the density in the receiver (through the high pressure valve), and
- the pressure in the receiver increases the density in the gas cooler.

Probably, the list could be continued, as every element on the colour map has some meaning.

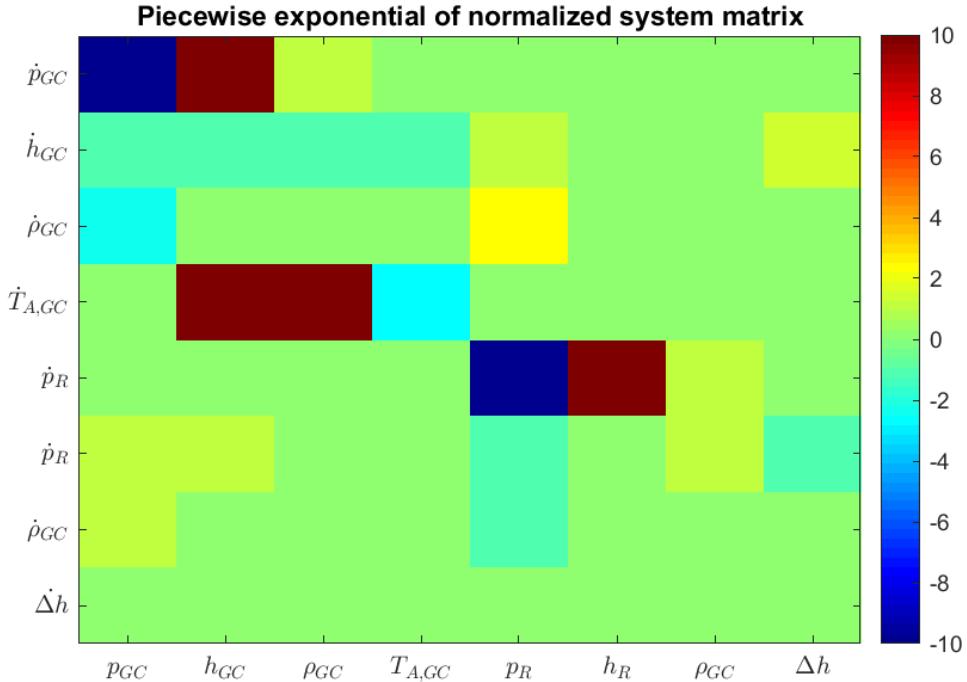


Figure 4.1: Illustration of system dynamics. Warm colour: *increase*, cold colour: *decrease*, green colour: *no effect*

4.1.2 Eigenvalue analysis

The poles, or eigenvalues of a linear system are denoted as $\text{eig}(\mathbf{A})$. Stability for a continuous linear system is described through the eigenvalues as

$$\begin{aligned} \text{Stable eigenvalue : } & \text{Real}\{\text{eig}(\mathbf{A})_i\} < 0, \\ \text{Unstable eigenvalue : } & \text{Real}\{\text{eig}(\mathbf{A})_i\} > 0, \\ \text{Integrator : } & \text{eig}(\mathbf{A})_i = 0. \end{aligned} \quad (4.4)$$

The eigenvalues of the system are unique and are not changed by linear transformations. If the system contains non-stable eigenvalues, in case of controllability, reference tracking can be achieved in closed loop by choosing the correct controller architecture [35].

In Figure 4.2, the piecewise exponentials of the imaginary and real parts of the eigenvalues of the system are colour coded, together with the transformation matrix from the states to the eigenstates. It can be seen that the system is dominated by two very stable poles (contributed by the pressure and enthalpy states), and there is a pair of oscillatory states, a slightly unstable state, and an integrator. The pressures contribute to the stability, since they show the highest sensitivity to energy misalignments in the system, counteracting with equalizing. The gas cooler enthalpy provides stability, due to enhancing the heat flow rate along the tubes. The fifth row of the transformation matrix describes the contribution to the real part of the oscillatory eigenstate-pair, and the sixth one belongs to the oscillations. It can be seen that the enthalpy of the receiver slightly contributes to the oscillatory behaviour, on the other hand, the enthalpy drop at the outlet of the gas cooler significantly damps the pair - this might be due to the rigid constraint of the valve on enthalpy (see Eq. 3.27). The enthalpy drop also plays the exclusive role for the last eigenstate, which is an integrator, since the state itself is an integrator.

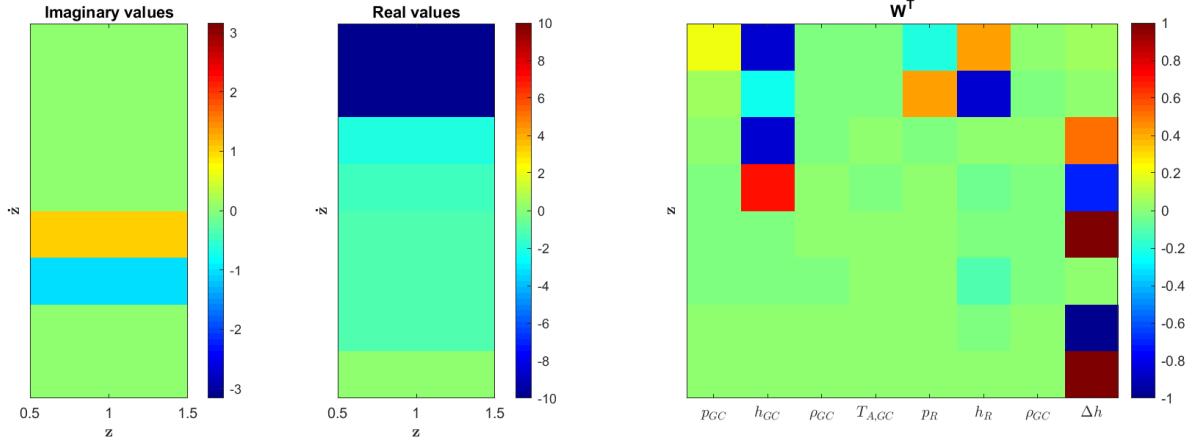


Figure 4.2: Piecewise exponential of the eigenvalues, and the transpose of the transformation matrix \mathbf{W} to the eigenstates. On the left, the colour coding refers to differential relation; on the right, it refers to linear relation

4.1.3 Controllability and observability analysis

Controllability and observability are achieved for linear time-invariant systems, if the controllability and observability matrices

$$\mathcal{C} = [\mathbf{A} \quad \mathbf{AB} \quad \dots \quad \mathbf{A}^{n-1}\mathbf{B}],$$

$$\mathcal{O} = \begin{bmatrix} \mathbf{A} \\ \mathbf{CA} \\ \dots \\ \mathbf{CA}^{n-1} \end{bmatrix} \quad (4.5)$$

have full rank for state number n , respectively. The staircase forms [46] transform the linear systems to separate the uncontrollable, or unobservable subsystems from the rest; hence, detectability and stabilizability can be analysed. For the transformations, Eq. 4.3 is used.

In the Appendix, Figure B.1 and Figure B.2 depict controllability, and the observability staircase forms, respectively. It can be seen that the system is entirely observable, however, a non-controllable integrator is found. This state is the enthalpy drop at the outlet of the gas cooler, Δh_{GC} , as it is a disturbance with zero derivative. Therefore the system is not stabilizable. However, by refining the gas cooler to have infinitely many cells, controllability could be shown. Nevertheless, lacking a high fidelity model prevents the possibility of designing a controller for it.

4.2 Linear time-varying system analysis

The linear time varying system is not so straightforward to analyse, as the time-invariant one. According to Lyapunov stability (see [42]), the stability analysis would need to be analysed through the range of system matrices, which might occur in the operation domain. However, due to the high model uncertainties, and the tedious calculations that would be needed, this is not investigated in depth. The literature (see [60] and [8]) generally considers the dependency of the system matrices according to time, while in our case, there is a strong nonlinear dependency on the states themselves, which can be regarded as feedback. Even if the methods are tried, the currently existing techniques for analysis are tedious. Time-invariant analysis may be considered to be well-motivated, since the operation point changes alter the parameters to a small extent. For supporting this, experiments were made with the linear time invariant (LTI) system based estimator to compete the time-varying one in Chapter 6, and although the results were worse, the LTI estimator was sufficient to determine

the states with decaying error signals.

With regard to controllability and observability, there is even less material providing black and white answers. This question is even more challenging, since the time evolution of the matrices is supposed to be predictable [57]. The original tool to consider controllability and observability, Grammians ([21]), are defined as being dependent on the system matrices. Frih et al. [16] use Bond graphs for controllability and observability analysis. Similarly to the stability analysis part, best said, if the system is considered to be sufficiently slow for control and observation with gain-scheduling.

Chapter 5

The Plants

In this chapter, the simulation and the real physical system are described.

5.1 Description of Modelica model

Two competitors in the market for simulations of this area are Matlab SimScape and TIL Suite library (click), built on Modelica language. The latter is a shallow object oriented library, which uses Modelica language to define the physics of a thermodynamical system, and uses the solvers of the IDE called Dymola (click). The solver is DASSL [47], which is a differential algebraic implicit solver. This means that it solves the equation of the form $0 = f(t, x, \dot{x})$, where x is the variable, by optimization. Since the initial conditions are usually not perfect, it starts from a local minimum, however, the controllers within the system, and the proper error handling (for example the treatment of extreme values) of the TIL library lets it to converge to normal operation after a while. Hence the initial ≈ 500 second period is not considered as valid simulation result.

In the earlier work of the author (see Appendix A), more reliable initial transients were provided, since the model used explicit solving method. However, explicit solver becomes very slow for the nonlinear differential equations, which are not stiff and use look-up tables, and the developed object oriented environment is more error prone in case of changes, than the TIL library one. Therefore, the simulation is carried out on a model built up in the commercial software, transported to Simulink, but for estimator and control design, the methodology for system description in the special course is used. The Simulink transported solver is solved by CVODE [14] solver in Matlab Simulink, which is a stiff/non-stiff differential algebraic solver. It has an implicit solver inbuilt for nonlinear subsystems. Note that this is a slower solver than DASSL, but according to experiments, it provides almost identical results.

The Modelica model, visualised on Figure 5.1 includes the system described in Chapter 5. Yellow colour indicates air, green indicates carbon-dioxide refrigerant, red indicates heat flow (in this latter physical domain, only heat transfer is considered). It uses 10 cells to finely model the heat transfer in the gas cooler. The compressor groups, and the two fans are modelled as one. The model has been adjusted in parameters to a real physical back-end system in the industry, and validated to have the same important characteristics as a real physical plant, for example equal heat transfer [15]. The evaporators are roughly modelled; the opening of them is simulated with an open system with a given mass flow rate, which induces the heat flow rate. This gives the capacitance-like behaviour, which is necessary to model gas loop in the operation domain. Due to the operation domain stated earlier, the heat recovery unit is constantly bypassed, and the gas cooler bypass valve is not used. The circles along flow lines (denoted by $\frac{dp}{dt}$) indicate pressure state flags, with which the user can denote, how many volumes are coexistent on different pressure levels. The small rectangles (denoted by Δp) along the flow lines ensure negligible pressure drops, which

ease the task of the optimizers of implicit solvers. The small circles branching from the flow lines are measurements, and they indicate, which physical quantity they measure (sh means superheat, which is analogue to enthalpy in our case). The physical dimensions of the system are picked based on the earlier mentioned existing plant [15], with the help of data sheets, and the same ones are used for the model established in Chapter 3.

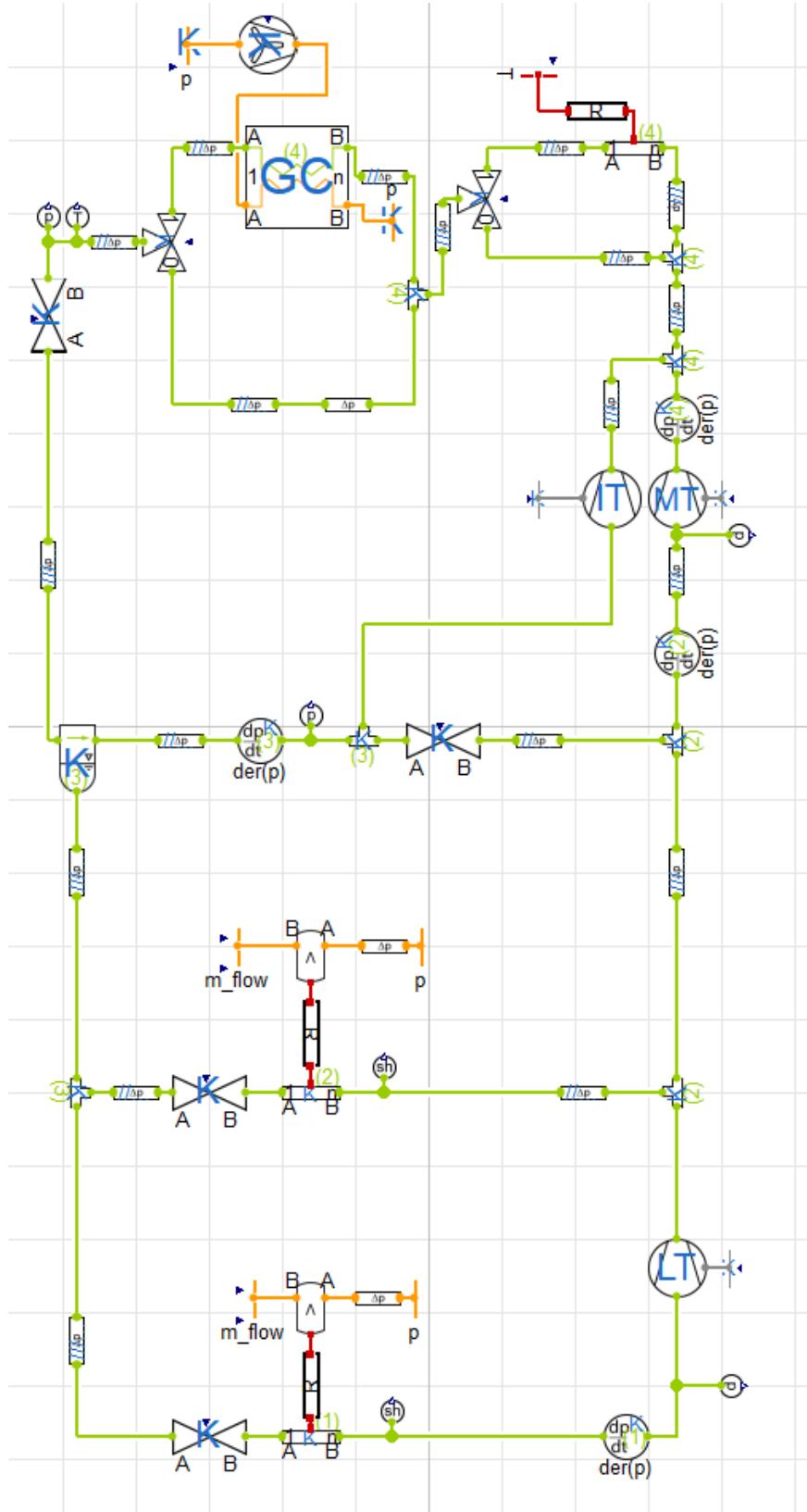


Figure 5.1: The Modelica model. Green: CO_2 . Yellow: gas. Red: heat transfer

The block is transported into Matlab Simulink as a functional mock-up unit (FMU, [7]). It is solved there with CVODE solver. Then, the Simulink solver itself can become a fixed-step explicit solver, as everything is defined in the discrete time domain, except the physical system. To ensure reasonable speed, Euler-backward method is chosen. Note, that due to the inputs are defined

as steps, the CVODE solver is a bit slower, than it would be with continuous time domain controllers.

The system was designed with TIL library (click). It provides a shallow object oriented structure to connect the parts of a thermodynamical system. For property conversion calculations, it uses a low-programming level COOLPROP [4] library.

The controllers acting on the system are distributed. They were tuned as continuous controllers in the Modelica model, then converted to discrete ones with 1 second sampling time. This sample time is chosen, since it is much faster than the dominant (thermodynamical) time constants of the system, which are about hundreds of seconds.

The reference values for the controllers, and the boundary conditions are presented in the Appendix, in Table D.1. The choice of pressure and temperature reference points for the outlet of the gas cooler are chosen as motivated in [54] and [50]. The temperature reference is some degrees higher than the ambient, or it lies with about 5 degrees superheat on the left of the saturation curve. It is necessary to keep some buffer there, since the system becomes unobservable, when the gas cooler outlet refrigerant becomes two-phase.

The input signals are low pass filtered with the time constant 100 seconds. The reason is that although the results are not significantly different, the simulator is running faster. The capacity ratio of the IT compressor is saturated at 0.7, since this was the only way of producing gas loop; it seems that the simulated refrigeration system is too robust.

The Simulink diagrams can be seen in Appendix C. The main parts of the simulation are the physical plant, the sensor emulator, the controller, the fault handling and the signal generator blocks (fault and ambient temperature). The sensor emulator block adds noises to the measurements and carries out property conversions. The controller sets the references, applies the distributed PI controllers, applies saturation and anti-windup (see [35]), then low pass filters the input signals to ease the job of the optimizer. The fault handling block consists of the fault handling logic, the RLS and the EM algorithms. Batch regularization may be applied. The data are saved out and then run offline on the estimator to speed up the procedure, enhancing the modular coding. Note that the modular pieces, like the PI controllers, the RLS, the fault detectors, the Kalman filter, and the set-point selector, are all programmed as Matlab objects, therefore they remember their parameters and states along the simulation, and only their functions are to be called. Calling external functions is solved by the INTERPRETED MATLAB FCN blocks.

The designed sinusoidal time evolution of the ambient temperature can be seen in Figure 5.2. It has decreasing dominant frequency, to fasten the training of the RLS algorithm at the beginning, but test data reminds the physical plant more; on the other hand, a smaller chirp (swept frequency) signal with increasing frequency is added, in order to test the robustness of the estimator. White noise is added as well, with the largest variance that the writer has experienced. The ambient temperature increases frequency until the edge of the scope of the project during the faulty operation period, therefore the range is entirely tested.

The fault itself can be seen in Figure 5.3. Two scenarios are tested: the one, which assumes a step fault and then switch back (in order to keep the simulation stable, a 10 second ramp is used for smoothing). The other way was to give a faulty measurement signal by averaging the true one with the ambient temperature, then switching back; this can be perceived as an insulation problem.

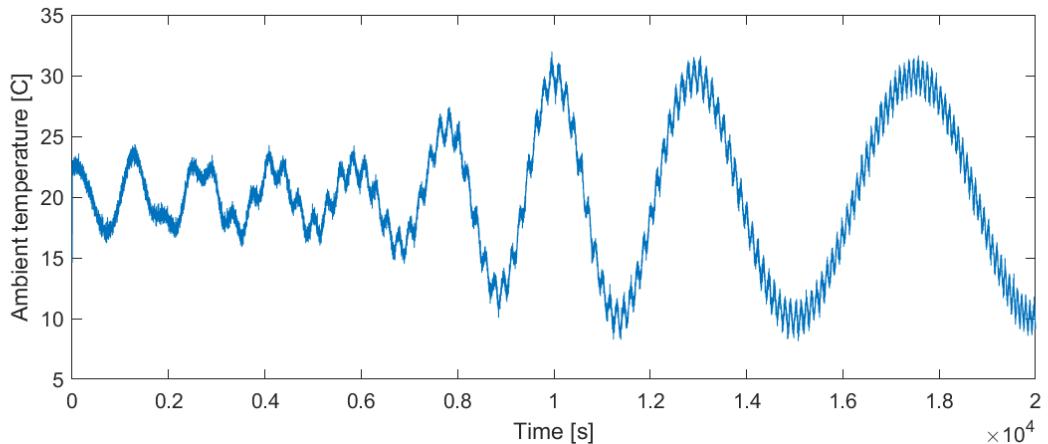


Figure 5.2: Simulated ambient temperature

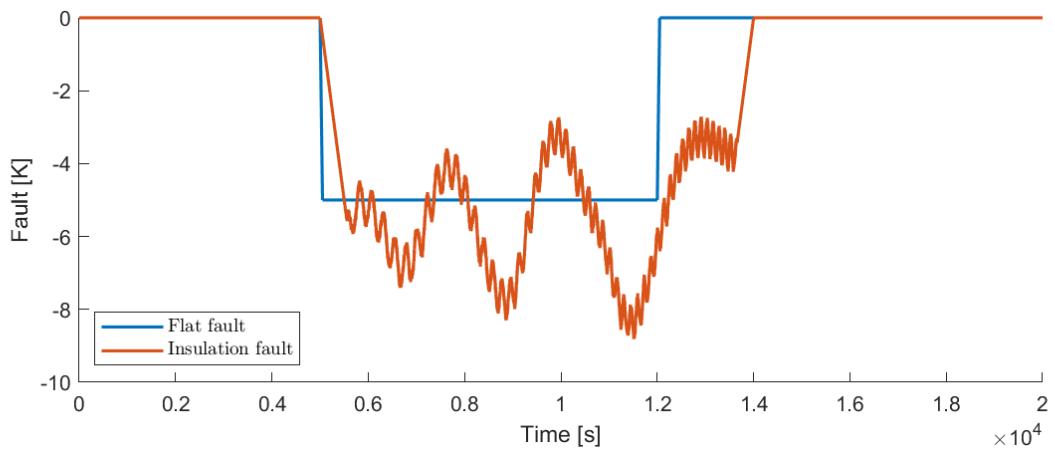


Figure 5.3: Simulated fault scenarios

5.2 Description of field data

The structure of the refrigeration system used for the field experiment is illustrated in Figure 5.4. In order to prevent liquid leaking to the suction side of the compressors, the refrigerant from the freezer is going through extra superheating, indirectly heat-exchanging with the receiver fluid. The sum of the outlet of the LT compressors and the middle temperature evaporators are led through a liquid buffer, to keep the liquid back. More importantly for the investigated application, bypass valve and heat recovery unit are not used. The physical parameters are picked based on the data sheets of the components [15]; it turned out that they are very similar to the ones in the simulation. Therefore, the parameters of the estimator model are not changed (small parameter uncertainties seem to have no effect on the estimations).

A photo of the field experiment apparatus can be seen in Appendix E, in Figure E.10.

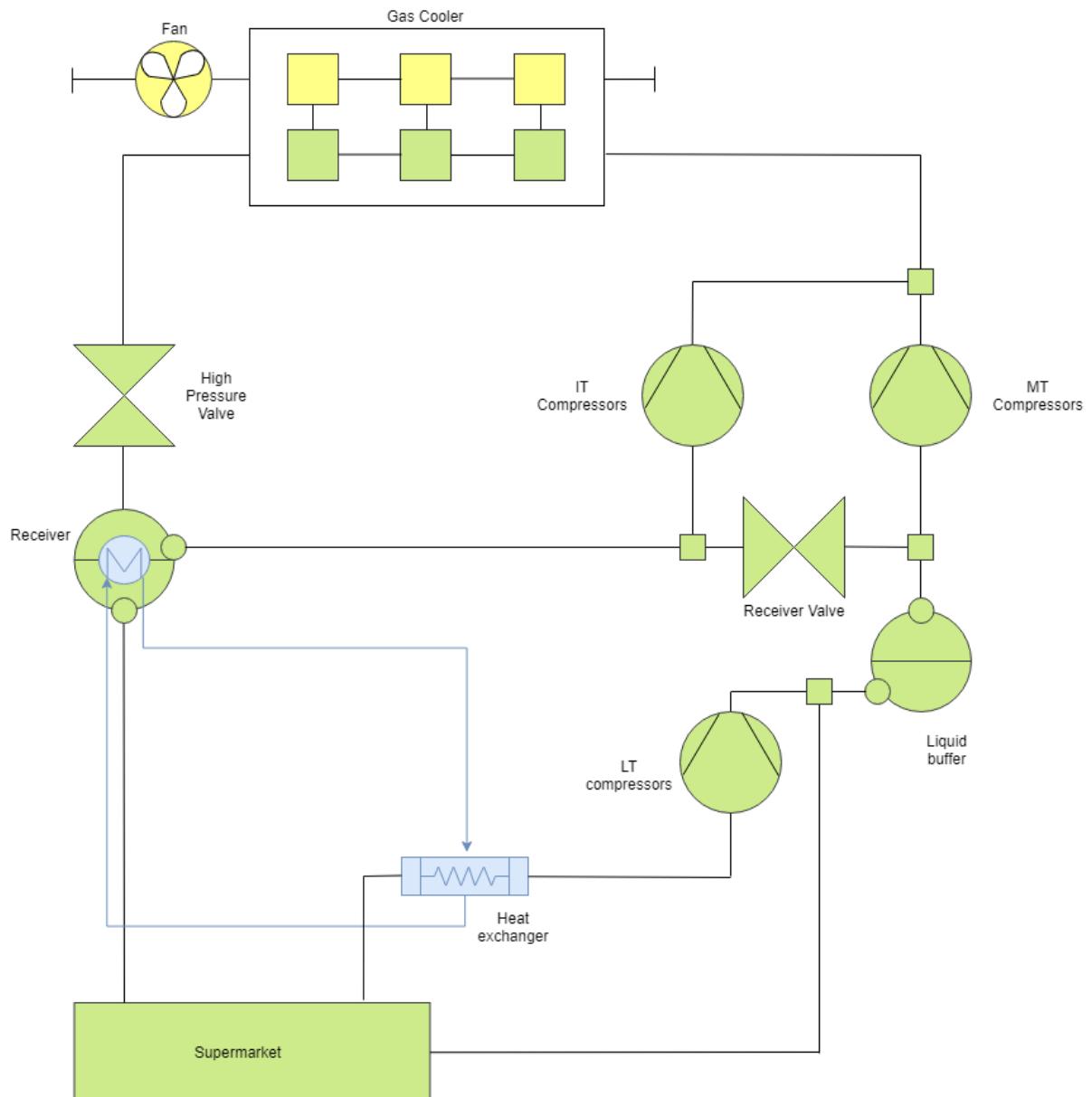


Figure 5.4: Refrigeration system for the field experiment

The ambient temperature was around 20 degrees, which is in the middle of the selected operation domain of the estimator and the fault detector. However, the fault did not have any effect, since there was a critically low level of refrigerant in the system (on the same plant, leakage detection algorithms were run), and due to the low density, the receiver was running under reference point. This means, that both the IT compressors, and the receiver valve were continuously closed, which prevents gas loop. To solve the problem, the ambient temperature measurement was reduced by five degrees, and since the optimal curve ([54] and [50]) then provided a reference point at a lower pressure, the density in the gas cooler and the receiver increased.

The gas loop handling of the industrial controller was then the following: it integrated the time of saturation of the IT compressors, or receiver valves, then it temporarily raised the reference point, to increase enthalpy-temperature gradient. This enhanced heat transfer, the system converged to normal operation, and the fault handling algorithm switched off. Later, since only the effects were cured, the procedure started from the beginning. The final result of this method is an oscillation, when most of the time the controller actually stays in gas loop.

The data needed to be treated. As it can be seen in Figure 5.5, the values, which were derived from the temperature of the outlet of the gas cooler with property conversion, have often performed jumps, or, they were undefined (this cannot be seen in the plot). Noise corrupted the signals, and being close to the saturation curve, for similar temperatures, the enthalpy values are often under the saturation curve, or on the right side of it. Not having better way to go around this latent heat problem (liquid level can only be measured in a buffer unit, where the phases clearly separate), the enthalpy, and the corresponding density values were located at the left side of the saturation curve.

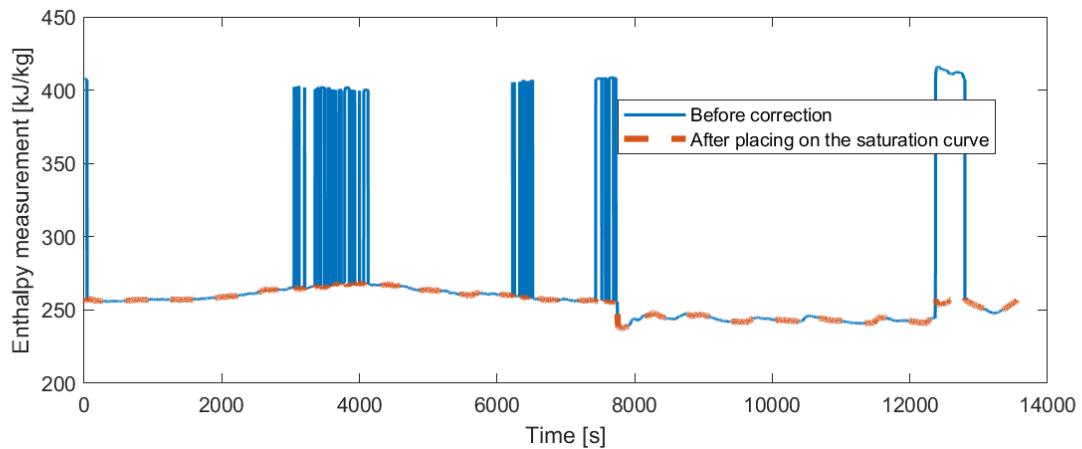


Figure 5.5: Treatment of gas cooler outlet temperature data

The fault scenario of the field data is depicted in Figure 5.6.

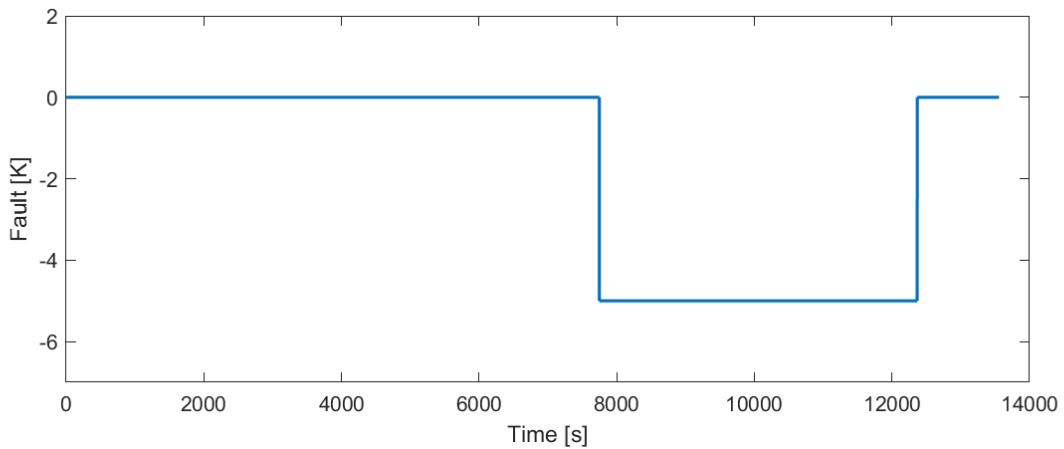


Figure 5.6: Fault for field data

The receiver liquid level compressor was not working, as it turned out from the measurements; it was showing the same (impossibly high) value all along the time. This value was indicating higher enthalpy, than the inlet side of the gas cooler, which is clearly wrong. Since the enthalpy measurement of the receiver is to be determined from the liquid level as well, and this measures an integrator state, the following was done. An enthalpy value was picked instead, which is lower than the largest possibly receivable enthalpy from the high pressure valve (assuming the 30 degrees ambient temperature operation point), and higher than the enthalpy level at the left side of the saturation curve.

Chapter 6

The State Estimator

In this section, the algorithms working in the two-fold estimator, then the motivations of the algorithms are described. The results are presented in Chapter 9.

6.1 Discretization

The estimator is a discretized estimator with sample time of $T_s = 1$ second. Therefore the continuous linear state space model is to be converted to its discrete version. It can be done by solving the equation

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \exp \begin{pmatrix} \mathbf{A}_{\text{cont}} & \mathbf{B}_{\text{cont}} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} T_s, \quad (6.1)$$
$$\mathbf{C} = \mathbf{C}_{\text{cont}},$$

where the index cont refers to the continuous equations.

6.2 Estimator architecture

The estimator architecture is presented in Algorithm 1. The abbreviation MA20 refers to the moving average of 20 samples, distribution \mathcal{Q} to the Gaussian Mixture Model distribution (see Section 8.3), the flag faultOp to the flag of the fault handling (it changes its value once from 0 to 1 during the algorithm). All the other variables will be addressed later along the documentation in more context.

There are two sampling times in the system: a 1 second one, which is for estimation of dynamics and detection of the fault. The slower sampling time, which is either 100 or 1000 seconds, sets the set-point (or operation-point, linearization-point) of the system. There are four main procedures done during the iterations: fault handling, a linear extended Kalman filter, a steady state filter, and set-point selection. The first one provides a fault detection method, which is independent of the other procedures, therefore it will only be described, where it becomes relevant: in Chapter 7 and Chapter 8. The rest is described in this chapter.

The general process consists of several steps. The result of the fault handling is given, the fault is either undetected (being present or not present), or is accommodated. The input and output deviations from the set-point values are calculated. Then the measurement and time update of the Kalman filter are done, doing one-step prediction ahead, based on innovation. If for some reason the filter does not end up being stable, the results are discarded, the values from the previous iteration are kept. In the steady state filter, those states, which are weakly observed, like the density and virtual air temperature of the gas cooler, are filtered based on their theoretical steady state values. There are physical extreme values in the system, which are enforced by soft constraints. During the set-point selection, the linearized analytic system matrices are calculated with substitution of

the inputs, the states, and using a look-up table. The linear filter is reset, and the stationary values are updated.

Algorithm 1: Two folded Kalman Filter loop

Sampling time: 1 second

procedure FAULT HANDLING (OPTIONAL)

RLS update: $\boldsymbol{\theta}, \boldsymbol{\varepsilon}, \mathbf{P}, \mathbf{K}_{\text{RLS}} \leftarrow \boldsymbol{\varphi}, \boldsymbol{\theta}, \mathbf{P}, \lambda$

EM update: $g, f, \mathcal{Q} \leftarrow \boldsymbol{\varepsilon}, h, \mathcal{Q}$ and earlier $\boldsymbol{\varepsilon}_M$ of window length M

if $\text{MA20}(f) > 0.9$ AND NOT faultOp **then** $\boldsymbol{\theta}_{\text{saved}} := \boldsymbol{\theta}$ from 500 samples before, **set** faultOp

if faultOp **then** $\boldsymbol{\theta} := \boldsymbol{\theta}_{\text{saved}}$, **override** h_{BPm} with \hat{h}_{BP}

procedure INNER FOLD: LINEAR KALMAN FILTER

Input-Output Deviation update: $\mathbf{u} := \mathbf{U} - \mathbf{U}_{\text{ss}}, \quad \mathbf{y} := \mathbf{Y} - \hat{\mathbf{Y}}_{\text{ss}}$

Measurement update: $\mathbf{x}_f, \mathbf{P}_f, \mathbf{K} \leftarrow \mathbf{y}, \hat{\mathbf{x}}, \mathbf{P}, \mathcal{M}$

Time update: $\hat{\mathbf{x}}, \mathbf{P}_f \leftarrow \hat{\mathbf{x}}_f, \mathbf{u}, \mathbf{P}_f, \mathcal{M}$

if $|\text{eig}(\mathbf{A} - \mathbf{KC})| \geq 1$ **then** disregard procedure

Sampling time: 100 or 1000 seconds

procedure OUTER FOLD: STEADY STATE FILTER

Time update: $\hat{\mathbf{X}} \leftarrow \mathbf{X}, \hat{\mathbf{x}}$

Measurement update: $\mathbf{X}_{f,\text{raw}} \leftarrow \hat{\mathbf{X}}, \mathbf{Y}$

Extreme value constraints: $\mathbf{X}_{f,c} \leftarrow \mathbf{X}_{f,\text{raw}}, \mathbf{X}_{\text{lim}}$

procedure SET POINT SELECTION

Look-up table and constant parameters: $\mathbf{p} \leftarrow \text{constants, look-up table}$

System matrix update: $\mathcal{M} \leftarrow \mathbf{X}_f, \mathbf{U}, \mathbf{p}$

reset $\hat{\mathbf{x}}$ to zero

Stationary values update: $\mathbf{U}_{\text{ss}} := \mathbf{U}, \quad \hat{\mathbf{Y}}_{\text{ss}} \leftarrow \mathbf{X}_f, \mathbf{U}, \mathcal{M}_{\text{nonlin}}$

6.3 The linear Kalman filter

The chosen linear estimator is a Kalman filter, since it is optimal for minimizing the estimation error variance (see [21]). The filter has a dynamic gain. On the other hand, stationary gain could be calculated, solving the infinite horizon estimation problem and the belonging discrete time algebraic Riccati equation. However, since the set-point is varying, the estimation process would have too large oscillation after set-point changes, if the uncertainties are not taken into account. Therefore the dynamic Kalman filter is chosen.

The Kalman filter algorithm is depicted in Eq. 6.2. Index k denotes the addressed time instant. Index $k|k-1$ stands for solving the optimization problem of the Kalman filter at time instant k , using the measurement values up to time instant $k-1$, while index $k|k$ denotes considering the newest measurement as well.

In the measurement update, the innovation (or residual) $\boldsymbol{\varepsilon}$ is amplified by the Kalman gain \mathbf{K} , to find a new (filtered) estimate $\hat{\mathbf{x}}_{k|k}$. The Kalman gain is calculated based on the 'ratio' of the state covariance matrix \mathbf{P} and the innovation covariance \mathbf{R}_ε . The latter is composed of the earlier value of the noiseless output covariance and the measurement covariance \mathbf{R} (the noises are assumed to be Gaussian, although, as described in Section 3.3.4, they are not). In the time update, the linear model $\mathcal{M}(\mathbf{A}, \mathbf{B}, \mathbf{C})$ and the filtered state with the input are used to find the prediction for the next time step $\hat{\mathbf{x}}_{k+1|k}$. Similarly, the state covariance goes through the dynamics estimation, and the input state covariance \mathbf{Q} is added to its value, which represents all the Gaussian zero mean

noise corrupting the states.

Measurement update:

$$\begin{aligned}
 \boldsymbol{\varepsilon}_k &= \mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1} \\
 \mathbf{R}_{\varepsilon,k} &= \mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R} \\
 \mathbf{K}_k &= \mathbf{P}_k\mathbf{C}^T\mathbf{R}_{\varepsilon,k}^{-1} \\
 \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\boldsymbol{\varepsilon}_k \\
 \mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{R}_{\varepsilon,k}\mathbf{K}_k^T
 \end{aligned} \tag{6.2}$$

Time update:

$$\begin{aligned}
 \hat{\mathbf{x}}_{k+1|k} &= \mathbf{A}\hat{\mathbf{x}}_{k|k} + \mathbf{B}\mathbf{u}_k \\
 \mathbf{P}_{k+1|k} &= \mathbf{A}\mathbf{P}_{k|k}\mathbf{A}^T + \mathbf{Q}
 \end{aligned}$$

In case of set-point changes, the estimator may become unstable temporarily. This means that the innovation is so contradictory, that the error covariance cannot be minimized, the estimation definitely does not converge to the true value. Therefore, in this case, the procedure is discarded and earlier values are used, until the problem disappears. Generally, this means one or two iterations.

Although it is assumed, that there is no covariance between the state and measurement noise distributions, some of the inputs are calculated based on measurements, through state conversions. (State conversions are done with the high programming level COOLPROP library.) Hence, the noise due to this correlation could be calculated. However, this is a seemingly negligible problem, caused by a nonlinear relation, furthermore, to calculate the noise states, several more computationally expensive equations would be introduced.

6.4 The steady-state filter

There are weakly observed states with autoregressive features in the Kalman filter, like the densities (integrators) and the virtual air temperature in the gas cooler (low pass filter). Due to these features, even if the process noise covariances are set well, sooner, or later, the filter forces the innovation to integrate in them. This results in slow drift. To solve this problem, they are 'pulled' to their analytical steady-state values, when the stationary states are updated. Since the filter is nonlinear, the weights are set manually. The state updates are the following:

Time update:

$$\mathbf{X}_{k|k-1} = \mathbf{X}_{k-1|k-1} + \hat{\mathbf{x}}_k \tag{6.3}$$

Steady-state update:

$$\mathbf{X}_{i,k|k} = \xi_i \mathbf{X}_{k|k-1} + (1 - \xi_i) \mathbf{X}_{\text{eq},i,k} \text{ for measured } i$$

Note that k is for another sampling time now. The last value of the linear Kalman filter prediction is added to the previously filtered stationary value $\mathbf{X}_{k-1|k-1}$ to create a prediction up to the current time instant. Then, in the measurement update, the weight $\xi_i \in (0, 1)$ is used for every measured state denoted by index i , and the equilibrium assumption is denoted by eq.

The steady state values are converted from the measured states for the densities. The virtual air temperature is given by the steady state version of Eq. 3.14; in other words, the state value is discarded. The value of ξ_i is 0.95 for the densities, and 0.5 for the virtual air temperature.

6.5 Extreme values and soft constraints

Sometimes the stationary states take up physically impossible values. This can be smaller due to inaccuracies, or larger after fault occurrence, and it results in large oscillations for several thousands of seconds. Therefore soft constraints were introduced. Since nonlinear values are to be constrained, stability conditions could possibly be found by the Karush–Kuhn–Tucker (KKT) conditions (see [28]), when numerical methods are considered. However, for this estimator, a closed form solution is addressed. The basic motivation is to make a compromise between the original value, and the constraint.

The constrainer can be seen in Figure 6.1. It lets the limit value to be exceeded to some extent at the beginning, then it gradually makes the extreme value to return to the limit value, finally lets it saturate there. The equation is the following, where index c denotes constrained, parameter $\alpha \in (0, 1)$ is the weight of the constraint, and parameter β is the maximum deviation from the limit value. A minimum constrainer can be easily created by inverting the values.

$$X_c = \begin{cases} X_{\text{raw}} & X_{\text{raw}} \leq X_{\lim} \\ \max(X_{\lim}, (1 - \alpha)X_{\text{raw}} + \alpha X_{\lim}) & X_{\text{raw}} > X_{\lim} \end{cases} \quad (6.4)$$

$$\alpha = \frac{1}{2\beta}(X_{\text{raw}} - X_{\lim})$$

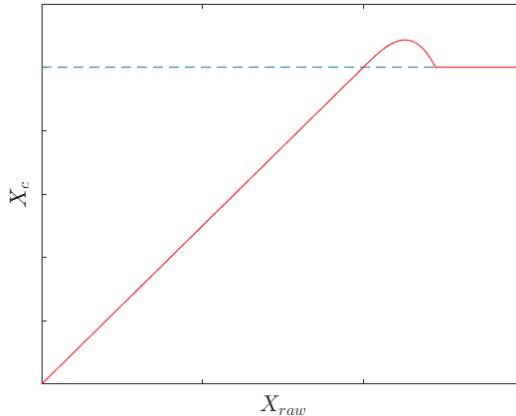


Figure 6.1: The constrainer function. The dashed line denotes the limit value

The following states were constrained:

- The densities cannot fall below zero.
- The virtual air temperature of the gas cooler cannot be lower, than the ambient temperature, since normally the surrounding ambient space of the gas cooler is supposed to receive the heat flow rate.
- The receiver enthalpy level is supposed to be bounded by the lower saturation line, and the enthalpy at the inlet.

One particular behaviour is observed for the sampled constrainer, when faulty measurements are used for estimation. Due to the lack of supervision between the sampling instants, the linear estimator still tries to converge to the (incorrect) measurement values. However, oscillations are avoided by occasional supervision. The phenomenon will be presented with the other estimator results, in Section 9.3.

6.6 Set point selection

The set point is selected by substituting \mathbf{X}_{ss} and \mathbf{U}_{ss} into the parameter matrices for the linear differential equations, resulting from the equations in Section 3.4.3, and the look-up table values from Section 3.4.5. The look-up table values are found by reading the pressure and enthalpy state estimates. For computational efficiency, the symbolic equations being declared in Matlab are converted to functions, and the constant system parameters (for example volumes) are substituted during the initialization of the set-point selector. However, the discretization with Eq. 6.1 increases running time significantly.

The stationary values are updated, and the linear state prediction is reset to zero. To keep the model independent of the measurements, the output is estimated with the nonlinear equation

$$\hat{\mathbf{Y}}_{ss} = \mathbf{g}(\mathbf{X}_f, \mathbf{U}), \quad (6.5)$$

where g is the output equation from the nonlinear model $\mathcal{M}_{\text{nonlin}}$. This way of estimation does not provide convergence, but neither would the method of adding the linear outputs to the last measurement value. Furthermore, the latter method would need a bang-bang estimator [35] to avoid steady-state errors. These steady-state errors were experienced during experiments.

The stationary values are actually not stationary, but it is assumed that the dynamics of thermodynamics are sufficiently slow (meaning low derivatives), for letting the linear Kalman filter innovation to encompass the effect of neglecting the terms $\mathbf{f}(\mathbf{X}_{ss}, \mathbf{U}_{ss})$ (in Eq. 4.1) on the estimation.

6.7 Tuning of the estimator

The biggest challenge with the estimator is the tuning. The measurement noise covariances are measured, but there were eight states, for which process noise is to be found, without any analytical knowledge. In the end, it is a trial and error process.

In this paragraph, the tuning is considered for the continuous input state covariance matrix \mathbf{Q} . Covariance is now interpreted as weights for tuning the estimator. The following observations were made. The measured values shall receive smaller weights, since they are naturally close to the measurement values. The less the states are observable, the less weight they should receive (for example densities), else all the innovation is incorrectly plugged into these states. The more derived a state is, the more error it accumulates due to parameter uncertainties, hence it is advised to set larger weights for states like pressures.

Discretizing the input state covariance matrix with respect to time has been tried. Frequently updating the matrix was slow, even while using Cholesky factorization. Furthermore, the estimation error became unstable, due to the parameter uncertainties. This is because the discretization considers the noise propagation through the states along the sampling time instants. The equation used for deducing the discretized version of the matrix is

$$\begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \mathbf{0} & \mathbf{A} \end{pmatrix} = \exp \begin{pmatrix} -\mathbf{A}_{\text{cont}}^T & \mathbf{Q}_{\text{cont}} \\ \mathbf{0} & \mathbf{A}_{\text{cont}} \end{pmatrix} T_s, \quad (6.6)$$

$$\mathbf{Q} = \mathbf{A}^T \Phi_{12}.$$

Due to the problems mentioned above, the matrix is simplified as

$$\mathbf{Q} \approx \mathbf{Q}_{\text{cont}}. \quad (6.7)$$

6.8 Residuals and detectability

The parameters and the state estimations are inter-dependent, and if the estimates diverge too much from the true values, the estimator gets stuck in local minima, in case of unfortunate parameter-estimation combinations. The estimator stays close to the linearization point, if the estimator dynamics are significantly faster than the fastest dynamics in the system. If the pressure dynamics are not considered (as they are modelled for control design, but they are low pass filtered due to their speed), these dynamics are described by the mass balance equation (Eq. 3.4). The mass balance equation has the same dynamics as the actuator dynamics, which have about an order of magnitude higher time constants (see for example Figure 9.10), than the sampling time (1 second). Therefore large Kalman gain is used, which suppresses state, and hence innovation covariance.

The innovation signals, or so called residual signals, and the trace of the covariance matrix can be seen in Figure 6.2. The means of the residuals do not change in case of the fault, just the frequency content of them. However, the changing frequency of the ambient temperature has the same effect. The ambient temperature causes residual correlation, as it changes the set-point for the linear estimator, keeping it in transient phase. Additionally, the fault seems to just slightly increase the residual correlations caused by the operation point changes, therefore it 'hides'. Since fault detectors usually work more robustly, if the mean (and not the spectrum) of the signal is investigated, a new method is developed for detectability in Chapter 7.

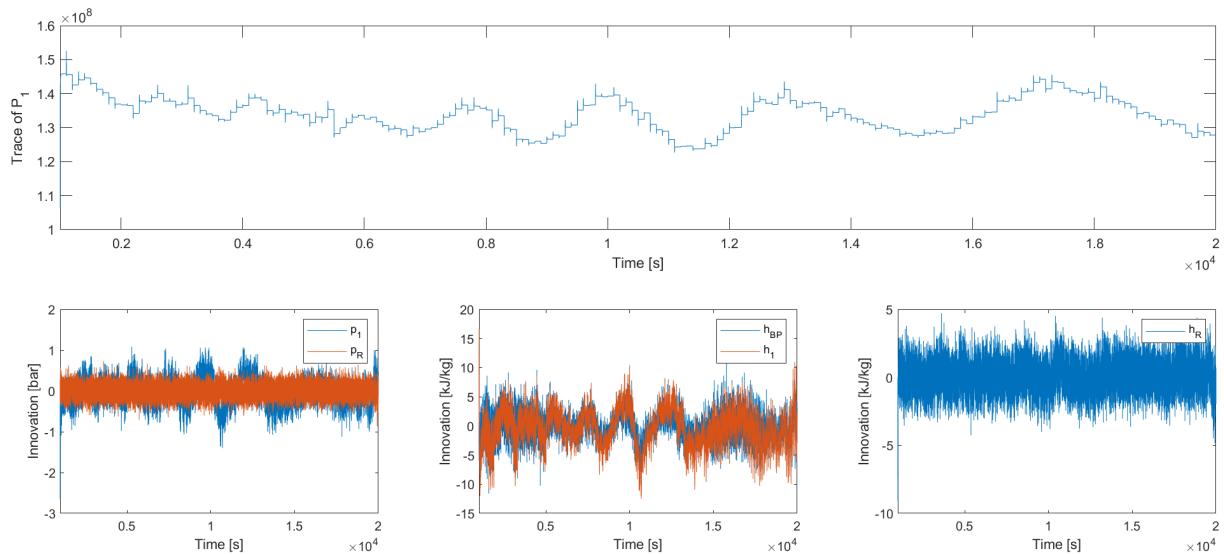


Figure 6.2: Innovation signals (down) and trace of covariance matrix (up) for the simulation signal, with no fault handling and insulation fault

One consideration may be to use the innovation of the outer fold. However, for set-point selection, low sampling frequency is desired, hence it had not been investigated in depth.

Chapter 7

Parameter Estimation and Residuals

Chapter 3 brings up the main challenges of the detection problem.

- The signals to be investigated are dynamical, hence steady-state correlations may be slow for detection.
- Slightly non-Gaussian signals are corrupting the measurements.
- The changes of the operation point have the same way of influencing the residuals of the linear filter, as the fault (see Section 6.8).

These are to be addressed and solved in this chapter, resulting in a residual signal with strong detectability. This is generated by the Recursive linear Least Squares (RLS) [13] algorithm, which is a computationally predictable and cheap algorithm.

7.1 Physical background

The basic idea is that the steady state energy balance should hold in the gas cooler, with some error range due to the transients. This error range should be bounded as the ambient temperature changes. Therefore, a coloured residual signal is expected from this algorithm. This behaviour should be captured by settling parameters after a while in normal operation. Then, when the fault occurs, the residuals should diverge temporarily from zero again. The energy balance equation (Eq. 3.5) is considered in the following steady-state form,

$$\begin{aligned}\dot{Q}_{\text{GC}} &= -\dot{m}_{\text{HR}}(h_{\text{HR}} - h_{\text{BP}}) = -\dot{m}_{\text{HV}}(h_{\text{HR}} - h_{\text{BP}}) = -(T_i - T_{\text{A},i})\rho_{\text{A},i-1}c_p \dot{V}_{\text{A}}, \\ \dot{m}_{\text{HR}} &= \dot{m}_{\text{MT}} + \dot{m}_{\text{IT}}, \\ \dot{m}_{\text{MT}} &= \rho_{\text{MT,in}} V_{\text{MT}} f_{\text{max}} \text{CR}_{\text{MT}}.\end{aligned}\tag{7.1}$$

Since the energy flows are weighted by parameters in the RLS algorithm, some flexibility is added. This data-driven technique results in a more adaptive algorithm, than the conservative model-based estimator in Chapter 6.

In order to find a structure, which represents energy balance sufficiently, all the energy flows of the gas cooler are considered, together with the input enthalpy, since these values are expected to be in a linear relation with the outlet enthalpy and the heat flow rate. Only measurements and inputs are considered, so that the estimator and the parameter estimation do not become inter-dependent. The distributed heat flow rate is difficult to be described by linear expressions; for simplicity, the volumetric flow rate and the temperature difference between the inlet temperatures (of the ambient air and the refrigerant) are separated into two separate regressors. The outputs \mathbf{y} are independently estimated, however, they can be easily calculated in matrix-vector form, using the algorithm is Section 7.2. The constant, or relatively constant values are dropped, since the

RLS algorithm will still solve an equivalent problem. The variable $\boldsymbol{\varphi}$ is the regressor (vector) and $\boldsymbol{\theta}$ is the parameter vector. In Eq. 7.2, the accounted variables are the capacity ratios CR_i , and the temperature difference of the media inflows $T_{HR} - T_{A0}$, the inlet enthalpy h_{HR} , finally, the pressure difference $p_{GC} - p_R$, giving the linear regressor $\boldsymbol{\varphi}$. All these variables are known directly or indirectly, from input signals and measurements.

$$\begin{aligned}\hat{\mathbf{y}} &= \boldsymbol{\varphi}^T \boldsymbol{\theta} \\ \mathbf{y} &= [\dot{Q}_{GC} \ h_{BP}] , \quad \text{where} \\ \dot{Q}_{GC} &= -\dot{m}_{HR}(h_{HR} - h_{BP}), \\ \boldsymbol{\varphi} &= \begin{bmatrix} 1 \\ h_{HR} \\ T_{HR} - T_{A0} \\ CR_A \\ \dot{m}_{MT}h_{HR} \\ \dot{m}_{IT}h_{HR} \\ CR_{HV}\sqrt{p_{GC} - p_R} \end{bmatrix} \end{aligned} \tag{7.2}$$

A dynamic model was considered as well, which was using earlier values (for one or two iterations) of the output signal in the regressor. However, that model encompassed the residuals, abolishing detectability. Note that the LTV estimator was not working for the same reason: the fault is 'hidden' enough to be accounted for changing dynamics.

7.2 Recursive least squares algorithm

The RLS algorithm is described in Eq. 7.3. The residual $\boldsymbol{\varepsilon}$ is given as the difference between the output and its estimation. Then the Schur-complement s is given as the denominator of the feedback of the covariance \mathbf{P} . The Kalman gain \mathbf{K} is then set to minimize the covariance. The parameter vector and its covariance are updated afterwards. The covariance is decreased by the update, but more information content of the regressors slow the decay. In order to ensure continuous learning, the forgetting factor λ is introduced [27], which is between 0 and 1 (chosen as 1-1e-4 in order to keep sufficient stability and learning at the same time). The smaller this value, the more the RLS algorithm forgets. The forgetting factor can be viewed as a way of regularization, which will be described in Section 7.3.

$$\begin{aligned}\boldsymbol{\varepsilon}_{k+1} &= \mathbf{y}_{k+1} - \boldsymbol{\varphi}_{k+1}^T \hat{\boldsymbol{\theta}}_k \\ s_{k+1} &= \lambda + \boldsymbol{\varphi}_{k+1}^T \mathbf{P}_k \boldsymbol{\varphi}_{k+1} \\ \mathbf{K}_{k+1} &= \frac{\mathbf{P}_k \boldsymbol{\varphi}_{k+1}}{s_{k+1}} \\ \hat{\boldsymbol{\theta}}_{k+1} &= \hat{\boldsymbol{\theta}}_k + \mathbf{K}_{k+1} \boldsymbol{\varepsilon}_{k+1} \\ \mathbf{P}_{k+1} &= \frac{\mathbf{P}_k - \mathbf{K}_{k+1} s_{k+1} \mathbf{K}_{k+1}^T}{\lambda} \end{aligned} \tag{7.3}$$

Note that although the calculations for the two residuals are done simultaneously, they are independent from each other.

The RLS algorithm is not switched on, until the first transients of the simulation ended. This is in order to avoid non-descriptive operation points, which are due to the initialization of the implicit solver. This helps the parameter vector to settle in the right part of the parameter space at the beginning.

7.3 Regularization and robustness

The parameter estimation problem turned out to be likely to overfit the training operation point. Since overfitting ends up incorrectly generalizing the local characteristics, robustness and applicability decays. Therefore, different regularization methods were tried. A method is successful, if it decreases the test condition sensitivity.

The successful methods were the following.

- The frequencies of the main input to the refrigeration system, the ambient temperature, are continuously changing. The larger amplitude signal has a decreasing, while the smaller amplitude signal has increasing frequency (see Figure 5.2).
- The amplitude of the ambient temperature is started at a smaller value, having the signal around the middle of the operation domain, to ease the optimization problem. Then, it was continuously increased. It was still increasing during the testing period, when fault handling was acting. This did not significantly deteriorate the results, which shows that the strategy is successful.
- The use of forgetting factor improves the results significantly. It is the equivalent to the weight regularization of the static least squares algorithm (see [5]). Therefore, according to bias-variance decomposition, the local variance increases, while the bias from the local output mean increases. This makes it possible to capture only the permanent behaviour of the data.

The unsuccessful methods were the following.

- Batch regularization was used. Earlier data from different linearization points were continuously mixed into the training data (the distances in time were around half and quarter of the period of the ambient temperature). Stability could only be provided by not changing the covariance matrix and setting the forgetting factor to 1, during the batch trainings. The results showed almost negligible improvements.
- White noise perturbation was superimposed on the parameters, having initially 10 % standard deviation of the values of the parameters, decreasing exponentially towards zero (settling in some periods). The results did not change. When the variance of the noise was increased, the estimation became unstable.

7.4 Delay calculations

In order to set an informative correspondence for the RLS algorithm, the causes and effects need to be connected along time dimension, in other words, delays need to be considered. This delay is possibly dependent on the linearization point, but this dependency is neglected. The delays were investigated by the cross-covariance function in Matlab, however, they reflected a range of correlations, which were dependent on the spectrum of the ambient temperature, since all the physical quantities are connected to this one through controller feedbacks. These connections result in correlations, which corrupt the cross-covariance functions.

Therefore the delays were approximated from the physical model, taking the low pass filter for the temperature of the air, and the enthalpy inlet of the gascooler in consideration. Smaller delay assumption is more reliable, as there are less side effects (as time passes, further correlations pop up due to the different interactions in the system). Furthermore, the value after the first time constant

is already more than 60 % of the steady state value, the delay is chosen as the time constant from Eq. 3.14 and Eq. 3.11,

$$\begin{aligned} T_{\text{delay},A} &:= \tau_A, \\ T_{\text{delay},\text{GC}} &:= \frac{\dot{m}}{\rho V_{\text{GC}}}. \end{aligned} \quad (7.4)$$

It is assumed that all the physical quantities that connect to the inlet of the refrigerant side have the delay $T_{\text{delay},\text{GC}} \approx 50$ seconds, and the ones connecting to the air side $T_{\text{delay},A} \approx 150$ seconds. The capacity ratio of the valve, which is at the outlet of the valve, is not delayed, as most of the change occurs at the inlets.

The value found for the time constant in Eq. 3.14 neglects some dominant dynamics, namely thermodynamical dynamics of the system. Figure 7.1 presents the way to estimate the time delay and time constant choice for the ambient temperature, which is the cross-covariance function. The first peak of the function indicates, at which lag the signals have most in common. Another way is to estimate the time constant with the Matlab function TFEST, which gave 170 seconds, being similar to the measured 125 seconds in the figure. Note that the delays are dependent on the frequency, for example, for the field data, no delay can be shown (however, at this slowness, a modelled delay value of 150 seconds does not have any effect on estimations).

This delay, or time constant is investigated at the outlet of the gas cooler. However, for the state estimator, the value of the virtual air temperature time constant is the matter of interest. Hence, it is supposed to be somewhat larger for the deeper points. Nevertheless, it seems that setting a time constant between 100 and 400 seconds gives about equally good estimations.

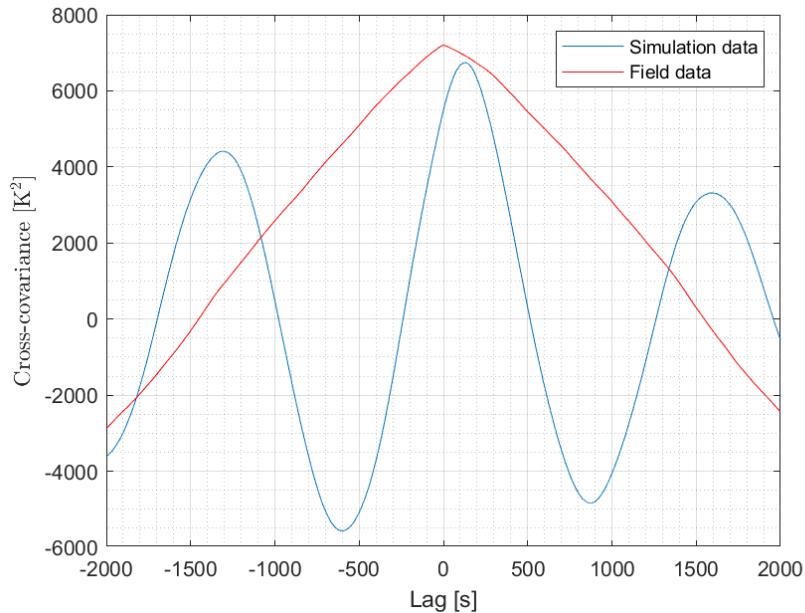


Figure 7.1: Cross covariance between the ambient temperature and the gas cooler outlet temperature, normal operation

7.5 The problem with information density

The residuals, and later, the fault estimations are very dependent on the quality of the signals. The capacity ratio CR_{IT} is included in the regressor vector, since it is an indirect input of the gas

cooler, and its saturation is used in the industry to detect this kind of fault. However, its saturation at the extreme values (0,1) means that the refrigeration system has already left its energy efficient stage; and it is often saturated during the parameter estimation procedure. This possibly results in the coefficients to be useless, therefore the covariance become large. The parameter covariances can be described with Eq. 7.5, where Σ is the covariance matrix. The standard deviations were normalized with the parameter values, and these relative covariances turn out to be occasionally extremely large. In case of an industrial application, robustness is required, therefore this problem needs special attention.

$$\Sigma_{\theta} = \Sigma_y (\varphi^T \varphi)^{-1} \quad (7.5)$$

In a special case, for the field data, the capacity ratio of the fans CR_A was saturated as well during normal operation, but reduced during faulty operation. This naturally causes an error in the fault estimation, which is linearly proportional to the deviation of the capacity ratio from 1. Hence, this value was overridden with the (earlier) value 1 in the field data case.

Alternatively, the information density problem could be improved with superposing perturbation signals with low crest factor (see [3]) on the input signals, for example, pseudo random binary sequence (PRBS) signal. However, this may face physical limitations during realization, due to that the COP values would decrease in case of large perturbation amplitudes, or the actuator resolution would not pass the information content sufficiently in case of small perturbation amplitudes.

7.6 Strong detectability

The detectability is weak, due to the nature of the RLS algorithm. This hides the fault in the long term, and corrupts the fault estimation. Therefore, the safety time of 500 seconds is chosen to sample the parameter vector θ_{saved} back in time, and the estimation $\hat{y} = \varphi^T \theta_{\text{saved}}$ is used to replace the faulty measurement.

7.7 Parameter identification for the state estimator

The two parameters s_0 and k from Eq. 3.14 are roughly estimated, based on the operation point. It was tried to actually identify their values, either independently, or derived from the parameters given by the RLS algorithm. The former assumes output, regressor and parameter vector

$$\begin{aligned} y &= \sigma_{\text{conv}} = \frac{\dot{Q}}{T_{\text{BP}} - T_{A,0}}, \\ \varphi^T &= [1 \quad \dot{V}_A], \\ \theta &= [s_0 \quad k], \end{aligned} \quad (7.6)$$

assuming that the temperature difference is constant along the gas cooler tube. This is a rough assumption as seen in Figure 3.7. The other method calculates the value from the parameters of Eq. 7.2 and Eq. 7.3, with the same assumption,

$$\begin{aligned} \hat{s}_0 &:= \frac{\dot{\hat{Q}} - \theta(2)\dot{V}_A}{T_{\text{BP}} - T_{A,0}}, \\ \hat{k} &:= \frac{\theta(2)}{T_{\text{BP}} - T_{A,0}}. \end{aligned} \quad (7.7)$$

This further assumes that the part of the convection content, which is not dependent on the volumetric flow rate, can be approximated by a constant. The results are almost identical for the two strategies, however, they both get stuck in a local minimum, where the value of k is a small

negative number, while s_0 has a big range along the operation domain. This indicates that the linear structure chosen is only a rough model of the true behaviour. For this reason, and in order to increase modularity, the results are not used to update the parameters of the estimator.

7.8 Results

The time evolution of the residuals are depicted together with the fault diagnosis in the following chapter. Both for the simulation and the field data, the faulty periods can be easily identified by naked eye (see Figure 9.1 and Figure 9.2).

Some of the detectors presented in Chapter 8 work better with white residuals. Hence, whiteness is investigated. Only the first residual is taken into consideration, which is explained in the same chapter.

The methods to describe residual whitening are described in Section 7.9. For the simulation data, the structure choice showed that the best whitening filter (or inverse filter) is a static filter. However, this is equivalent to repeating a reduced version of the RLS algorithm, therefore whitening is not applied. Figure 7.2 shows that the data is actually very close to being white and normal. Whiteness is statistically achieved if the data is under the confidence bounds (assuming a moving average process) in the sample autocorrelation function, or is achieved with 99 % confidence interval, if the line is bounded by the two straight lines on the cumulative periodogram. Normal probability distribution is investigated, since the process is a complicated nonlinear process, and according to central limit theorem [5], summing up different distributions, the result is converging to this kind of distribution. This means that complex industrial processes are most likely to have normally distributed signals. Sending these normal distributed signals through RLS algorithm will result in a normal distributed residual signal. Testing this, the probability distribution seems to be almost normal, since the probability plot and the histogram are fitting the theoretical lines well. (The data was tested for t-distribution and logistic distribution as well, giving worse fit than for normal.)

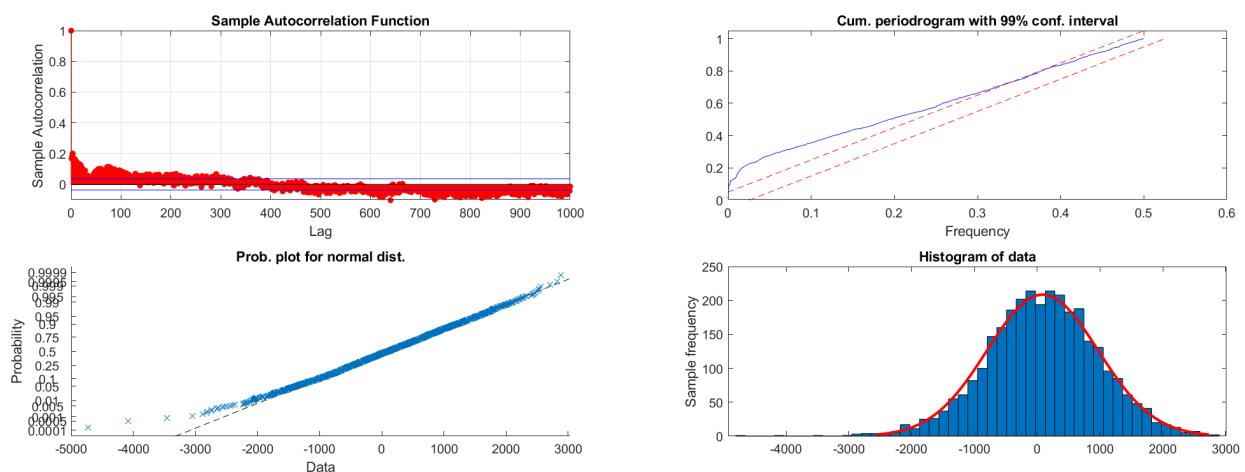


Figure 7.2: Residual statistics for simulation data. Statistics of whiteness (up) and fitting normal distribution (down)

The results were not so positive for the field data. Therefore, residual whitening is applied in Section 7.9.

7.9 Residual whitening

The whitening of the residuals is necessary for CUSUM and GLR, as it is described in Section 8. There are two ways to do that. One of them is that certain frequencies are passed or stopped, given that the spectrum of the signal is known. This has been done with a Butterworth filter [53]. However, the spectrum of the residual is dependent on the input signals, hence it changes and robustness cannot be guaranteed. Therefore another method is used, which proves to be less sensitive to these effects: parameter identification.

The basic concept is that, with the following ARMAX model [62], the residual $\varepsilon(t)$ can be viewed as the sum of an input $u(t)$ and Gaussian noise $n(t)$,

$$A(q)\varepsilon(t) = B(q)u(t) + C(q)n(t), \quad (7.8)$$

where the polynomials $A(q)$, $B(q)$ and $C(q)$ are dependent on the delay operator q , meaning that both moving averages and autoregressions are used. The whitening can be done by using the inverse filter, resulting from solving Eq. 7.8 for $n(t)$.

7.9.1 Input signal choice

If the front-end of the refrigeration system is assumed to work perfectly, only the ambient temperature and the input signal nonlinearities (saturation, discretized value, etc.) can influence the residuals. Since in normal operation, the latter does not have significant effects, only the ambient temperature is considered. Since, in the investigated model, all the capacity ratios, and hence physical quantities, are correlated with the ambient temperature through the controllers or physical phenomena, this is the best choice for a simplistic model.

7.9.2 Structure choice criteria

Too high structure order causes overfitting (see [5]), which prevents successful generalization for application purposes. Therefore more structure variations were tested. This means all combinations up to structure parameter order 10. The best structure model is identified based on several evaluation criteria, which are the following:

- F-test,
- Information criteria (Akaike Information Criterion, Bayesian Information Criterion, Final Prediction Error Criterion),
- Autocorrelation and cross-correlation of the noise with the input signal,
- Cumulative periodogram,
- Significance and consistency of identified poles and zeros,
- Conditions of covariance matrices.

In order to avoid overfitting of data sets, cross-validation is included (see [5]), the data is divided into two partitions (one is the train set and the other is the test set), and the results, which are the least worst for both test set partitions, are preferred. The two set combinations are called forward and backward directions. For every model order number, all the combinations of the orders n_A, n_B, n_C are investigated, and the best one is chosen for further evaluation.

The residual of the filter turns out to be white from model order 5 up to higher model orders, as the autocorrelation function, the cross-correlation with the input signal, and the cumulative periodogram showed. Since these results do not play role in the final selection, they are not depicted.

The information criteria [1] punish structure order d , using test data loss function W_N , where N is the number of the data points (see Eq. 7.9). They are chosen as the first criteria, since usually the most important dynamics are captured at the beginning of the first plateau after the first drop, if plotted against model order. The results can be seen in Figure 7.3. It seems that here, at model order 5 and 6, these criteria are minimized. It is interesting to note that the set choice significantly changes the validation results. Therefore, for future work, a better resolution of cross-validation is advised.

$$\begin{aligned} \text{AIC} &= (1 + \frac{2d}{N})W_N, \\ \text{BIC} &= (1 + \frac{\log Nd}{N})W_N, \\ \text{FPE} &= \frac{N + D}{N - d}W_N \approx (1 + \frac{2d}{N} + ...)W_N. \end{aligned} \quad (7.9)$$

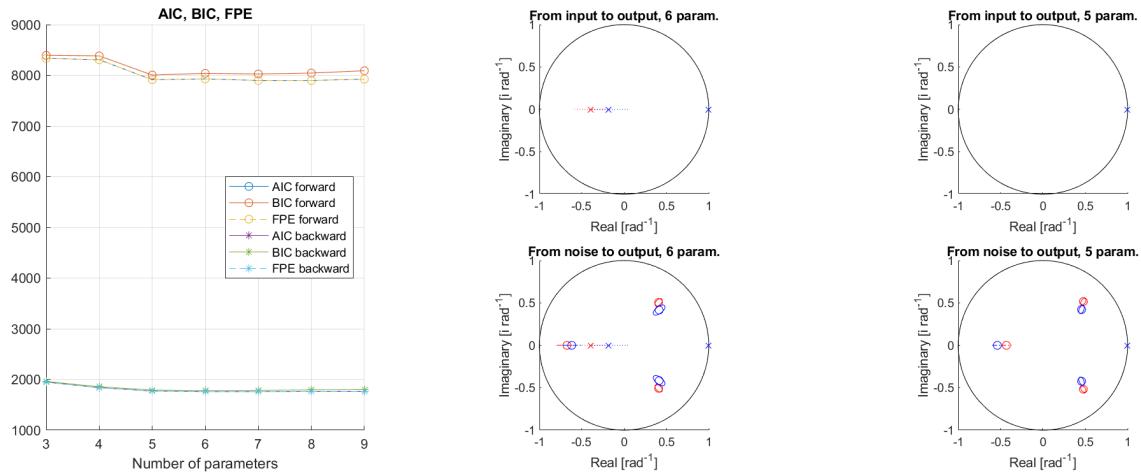


Figure 7.3: Statistical criteria for structure choice. Left: information criteria. Right: identified poles and zeros for two cross-validation sets, with ellipses indicating 95% confidence interval

On the right side of Figure 7.3, the significance and covariance of the poles and zeros are compared. Blue and red are indicating the two different set distributions (forward and backward). The pole, which is added by the 6th order model, is close to the middle of the unit circle, showing small significance. Furthermore, the covariances of the zeros increase for the higher model order as well.

Regarding the condition numbers of the covariance matrices, they can be seen in Table 7.1. These matrices have high condition numbers, if the models are over-parameterized. However, a natural increase is expected in case of increasing model parameters as well. It can be seen that model orders 5 and 6 seem to produce lower condition numbers.

Model order	Forward direction	Backward direction
8	2.14e9	5.74e11
7	1.47e9	1.3e11
6	2.04e6	5.12e6
5	1.02e6	2.00e6

Table 7.1: Condition number of covariance matrices of identified parameters

The F test [41] is defined as F-distribution evaluation of the zero hypothesis, that adding a parameter to a model explains the underlying model better:

$$\mathcal{H}_0 : \quad z = \frac{J_1 - J_2}{J_2} \frac{n - d_2}{d_2 - d_1} \in \mathcal{F}(d_2 - d_1, N - d_2), \quad (7.10)$$

where J is the loss function W_N divided by 2, index 2 denotes the new and index 1 denotes the old model. It is the following for the two identification directions:

$$\begin{aligned} \mathcal{F}_{\text{forward}} : & \left[\begin{array}{c|ccc} \text{from/to} & 4 & 5 & 6 \\ \hline 3 & 100 & 100 & 100 \\ 4 & 0 & 100 & 100 \\ 5 & 0 & 0 & 99 \end{array} \right] \\ \mathcal{F}_{\text{backward}} : & \left[\begin{array}{c|ccc} \text{from/to} & 4 & 5 & 6 \\ \hline 3 & 100 & 100 & 100 \\ 4 & 0 & 100 & 100 \\ 5 & 0 & 0 & 100 \end{array} \right] \end{aligned} \quad (7.11)$$

First, note that the test is only valid for nested model orders. This is fulfilled. Then, if the value of the test is smaller than 100, then it seems that there is no added information about the structure. Since the forward test fails for increasing the model order to 6, and the same results are provided by investigating the poles and zeros, model order 6 with the best set-up, $n_A = 1, n_B = 1, n_C = 3$ is chosen. This indicates an autoregressive filter for the input, and an autoregressive-moving average filter for the noise.

7.9.3 Results

The results are depicted in Figure 7.4. The residual signal becomes significantly whiter after whitening (still having some correlation at the low frequency range), but since the data was slightly nonlinear, this nonlinearity is amplified by the filtering, and the semi-normal probability distribution is lost. Increasing whiteness results in better detection results, as the faulty and normal operations are easier to differentiate.

The time evolution of the whitened residual compared to the original will be depicted in Figure 9.2, together with the fault detectors.

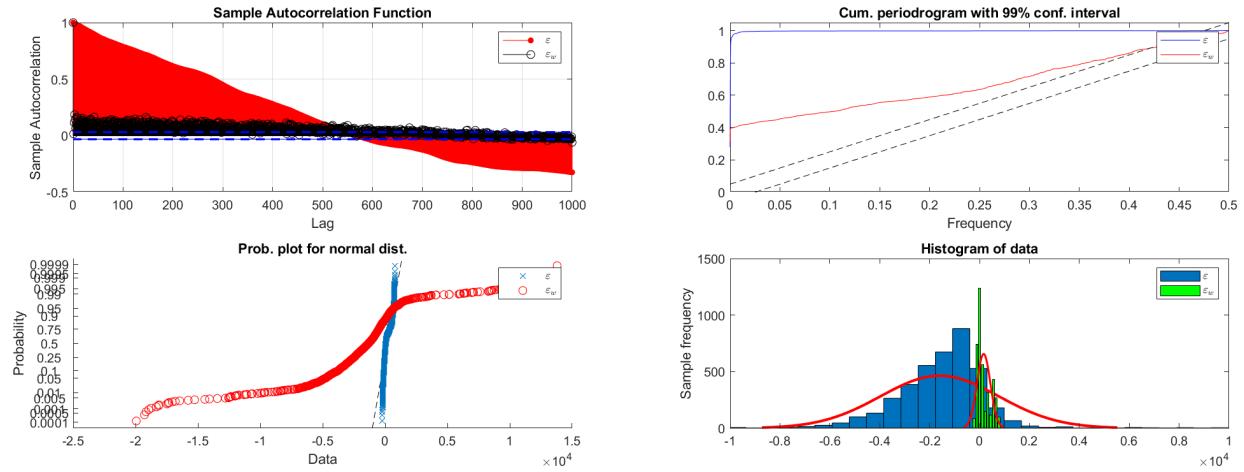


Figure 7.4: Residual statistics before and after whitening. Statistics of whiteness (up) and fitting normal distribution (down)

Chapter 8

Fault Detection and Operation

The residual signals are investigated with three different fault detectors. Two of them, the CUSUM and GLR [5], are applying autoregressive and moving average strategies, respectively. It is common for them, that they work better with Gaussian and white residuals. The third method, which is set up in this work based on Expectation Maximization (see [5]), only assumes Gaussianity, and its decision function ranges between 0 and 1, which makes it slightly harder to tune manually, but being easier to be interpreted.

In this chapter, first the fault detectors are discussed and compared, then fault handling is shown. Note that the comparison is not entirely fair, since the CUSUM and the GLR algorithm only use the residuals of the heat transfer in the gas cooler (as it is less dependent on the ambient temperature, than the outlet temperature, and produces faster transients in case of a fault), while the EM algorithm considers both dimensions of information. However, the multidimensional cases of the CUSUM and the GLR are either non-implementable or tedious to correctly implement, in case of these residuals.

One disadvantage of the residual signals in Figure 9.2 is that, when the fault is switched off, the transient effects take a long time to settle in normal operation. This tricks all the detectors to some extent.

The reasons why two residual signals are used, are the following. Both of them can be used independently, but the EM algorithm can easily enhance them simultaneously (which is not so true for the other methods, as it will be described). The heat flow rate residual diverges faster from the zero mean value in case of a fault; however, for fault estimation, the enthalpy residual gives better estimates. The former signal is chosen for the scalar fault detection methods.

Note that the variance of the residual signals is generally unknown and mostly unpredictable; it depends on the ambient temperature before that fault is detected. However, experiments show that it does not significantly change, therefore it is considered to be constant.

8.1 CUSUM detector

The CUSUM and the GLR algorithms rely on the Neyman-Pearson (see [45]) detection approach, using the log-likelihood ratio,

$$s(z) = \ln \frac{p_{\theta_1}(z)}{p_{\theta_0}(z)}, \quad (8.1)$$

where z is the observation, θ denotes the parameters of the distribution, index 1 the faulty, and

index 0 the normal operation. The expected values of this ratio hold the statistical property of being exclusively positive for the faulty operation, and negative for the normal,

$$\begin{aligned}\mathbb{E}_{\theta_0}(s) &= \int_{-\infty}^{\infty} s(z)p_{\theta_0}(z)dz < 0, \\ \mathbb{E}_{\theta_1}(s) &= \int_{-\infty}^{\infty} s(z)p_{\theta_1}(z)dz > 0.\end{aligned}\quad (8.2)$$

The distance between the cumulative sum $S(k)$, where k is the time index, and the minimum value of $S(k)$ becomes the decision function.

$$\begin{aligned}S(k) &= \sum_{i=1}^k s(z(i)) \\ g(k) &= S(k) - \min_{1 \leq j \leq k} S(j)\end{aligned}\quad (8.3)$$

Then, in case of Gaussian distribution, the following recursive decision function provides the quickest deviation from zero for a mean change towards the correct direction:

$$g(k) = \max \left(0, g(k-1) + \frac{\mu_1 - \mu_0}{\sigma^2} \left(z(k) - \frac{\mu_1 + \mu_0}{2} \right) \right) \quad (8.4)$$

The application of the algorithm is straightforward for multiple dimensions. However, the CUSUM algorithm needs better estimation of the faulty mean in order to show better performance. However, due to the curse of dimensionality [5], the variance increases quickly, as the number of dimensions are increased. This leads to both the normal and faulty likelihoods be equally very low, if they are not close enough to one of the pre-defined means. Therefore, the one-dimensional CUSUM is used.

Note that the theoretical Average Run Length values are not used for comparison, since the residual is either coloured or not sufficiently normal distributed as seen in Section 7.9.

The mean value was chosen as an inspected value after the simulations, as 10 kW. The threshold is chosen based on Eq. 8.4 to fire after 20 samples of step change,

$$h_{\text{CUSUM}} = 20 \frac{\mu_1^2}{2\sigma^2}. \quad (8.5)$$

8.2 GLR detector

For the generalized likelihood ratio algorithm, the mean is unknown. Then, the maximum likelihoods of the detection time and the faulty mean are evaluated in the cumulative sum (see Eq. 8.3). In order to bound computational time, a window length is considered, leading to the decision function

$$g(k) = \max_{k-M+1 \leq j \leq k} \max_{\theta_1} S_j^k(\theta_1). \quad (8.6)$$

Equating the derivative with zero, with respect to μ_1 , assuming Gaussian distributions, results in the decision function

$$g(k) = \frac{1}{2\sigma^2 M} \left[\sum_{i=k-M+1}^k (z(i) - \mu_0) \right], \quad (8.7)$$

where $M = 1000$ denotes the window length. Smaller window length resulted significantly worse hit rate (or detection rate).

The GLR application for vector case has one pitfall: the covariance matrix is supposed to be well known. This matrix changes very easily in the current situation for different simulation scenarios, therefore the vector case has not been tried. This problem can be viewed as the curse of dimensionality, just like described for the CUSUM case in Section 8.1.

Note that the theoretical probabilities of false alarm and missed detection are not used for comparison, since the residual is either coloured or not sufficiently normal distributed, as seen in Section 7.9.

The threshold is chosen based on the consideration that the decision function in Eq. 8.7 should step over it, when 20% of the samples in the window (therefore 20 out of 100) are at the mean value after a step change (the same mean value is used as for the CUSUM algorithm),

$$h_{\text{GLR}} = \frac{1}{2\sigma^2 M} (0.2\mu_1 M)^2. \quad (8.8)$$

8.3 Expectation Maximization detector

Since the residual vector is not white, alternative ways were searched, in order to set up a detector, which is computationally cheap, accurate, easy to interpret and robust to slight changes in the statistics of the residual vector. The Expectation Maximization algorithm [5] proved to be a good strategy. It fulfills all the criteria mentioned, except the first one, since it evaluates the exponential function 2 times of the window length in an iteration. However, the window can be set to small values (in this case the chosen value is 10), and then the detector still works robustly.

In order to understand the strategy, the probability distribution of the residual signal is to be seen as two clusters, which describe normal and faulty distributions. Since the best fittings are found for the normal distribution (see Section 7.9), it is taken as an assumption. Then, assuming that the distribution does not change in case of a fault, the general probability distribution can be described as a two-cluster Gaussian Mixture Model (GMM). The data is online recorded by a sliding window, then the expectation E and the maximization M steps are used to find the parameters and variables describing the distributions. This also gives the information, how much the data points are faulty or non-faulty; and when the average score weights more for the faulty distribution (scores more than half), fault is detected.

The posterior probabilities of the belonging to the clusters are called responsibilities (meaning how much each cluster are responsible for producing the data point), given the data \mathbf{x} . The clusters are indexed by k . The responsibilities $\gamma(z_k)$ add up to 1 for each data point, and piecewise belong both to the clusters, and to individual data points. The belonging z_k itself is the so called latent variable in this case, which is a binary value for each cluster: the data point either belongs to a certain cluster, or not. It is latent, because it is not observed, and considering a data point, it is 1 for only one cluster. The responsibility can also be viewed as the expected value of the latent variable. The marginal distribution over \mathbf{z} defines the mixing coefficient π_k , which helps to describe

the responsibility, such that

$$\begin{aligned}\pi_k &= p(z_k = 1) \\ \gamma(z_k) \equiv \mathbb{E}(z_k) &= p(z_k = 1|\mathbf{x}) = \frac{p(z_k = 1)p(\mathbf{x}|z_k = 1)}{\sum_{j=1}^K p(z_j = 1)p(\mathbf{x}|z_j = 1)} = \frac{\pi_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K \pi_j \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)}, \\ \sum_{k=1}^K \pi_k &= 1,\end{aligned}\quad (8.9)$$

where K is the number of clusters.

The EM algorithm can be viewed as minimizing the Kullback-Leibler (KL) divergence $\text{KL}(q||p)$. The KL divergence is a measure of the structural misalignment between two probability distributions. Since the true distribution p is unknown, the GMM is used as a distribution q . Then the log likelihood $\ln p(\mathbf{X}|\boldsymbol{\theta})$ is maximized for $\boldsymbol{\theta}$ by setting the distribution structure for $q(\mathbf{Z})$ with changing the responsibilities; this is the E step of EM. In this step, the structure is viewed as the true structure, giving virtually zero KL divergence. The model parameters $\boldsymbol{\theta}$ are set in the M step, to maximize the lower bound $\mathcal{L}(q, \boldsymbol{\theta})$. This step forces the parameters describe the data more, but allows the KL divergence to naturally diverge from zero. Note that even in this case, the log-likelihood will increase, since the step is formulated to increase the lower bound. The visualization of the process is described deeply in [5] and can be seen in Figure 8.1.

$$\begin{aligned}\ln p(\mathbf{X}|\boldsymbol{\theta}) &= \mathcal{L}(q, \boldsymbol{\theta}) + \text{KL}(q||p), \quad \text{where} \\ \mathcal{L}(q, \boldsymbol{\theta}) &= \sum_{\mathbf{Z}} q(\mathbf{Z}) \ln \frac{p(\mathbf{X}, \mathbf{Z}|\boldsymbol{\theta})}{q(\mathbf{Z})}, \\ \text{KL}(q||p) &= - \sum_{\mathbf{Z}} q(\mathbf{Z}) \ln \frac{p(\mathbf{Z}|\mathbf{X}, \boldsymbol{\theta})}{q(\mathbf{Z})}.\end{aligned}\quad (8.10)$$

For the GMM, the following log likelihood is maximized,

$$\ln p(\mathbf{X}|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{n=1}^N \ln \left\{ \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x}_n|\mathbf{x}_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \right\}. \quad (8.11)$$

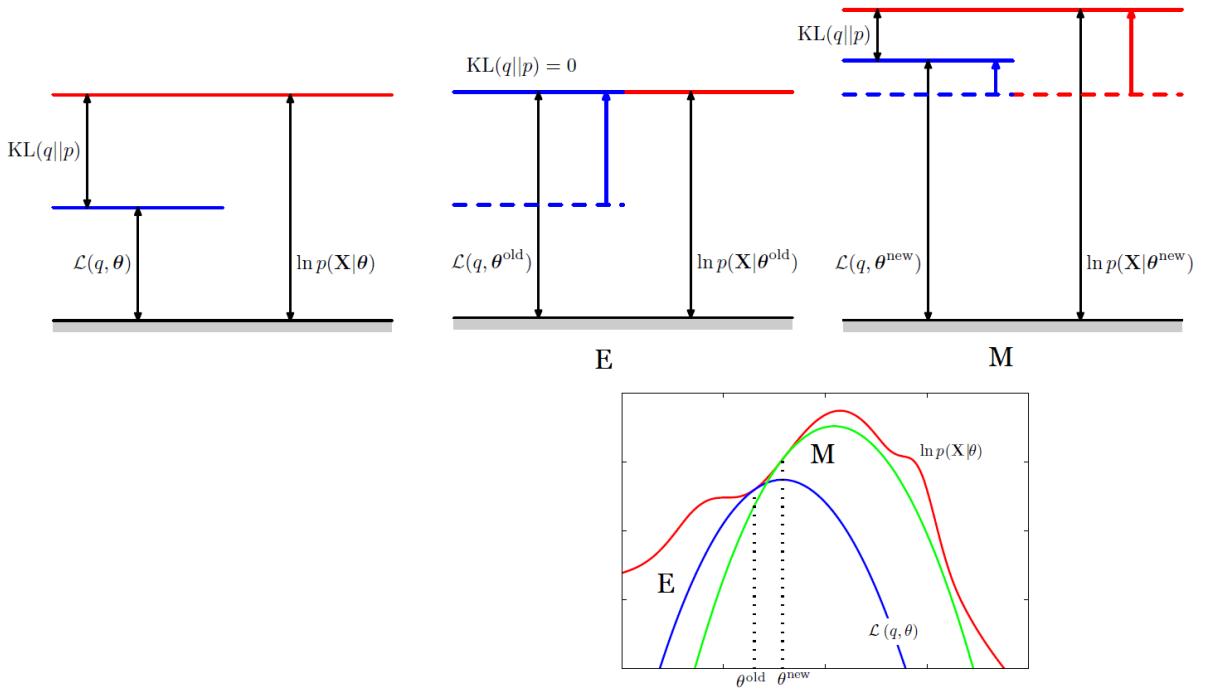


Figure 8.1: Visualization of the E and M steps [5]. The content of the log likelihood value originally, and after the steps (up); the visualization of the fitting (down)

There are some parameters, which are better left unmodified. First, the residual vector for normal operation has mostly zero mean. Second, having too slightly varying data, being close to their mean value, causes singularity for the variance estimation problem, as the estimate would decrease to a very small value. This will later give zero probability density scores to every further data points. It is also computationally expensive to update a more dimensional variance matrix. Therefore the residual variance vector for normal operation is found from experiments, and the one for faulty operation is set by tuning, as a scalar multiple of the normal operation one. The reason for that is that the faulty operation cluster is supposed to give higher probability density scores for the data points, which are far from zero, as illustrated in Figure 8.2, for one dimension.

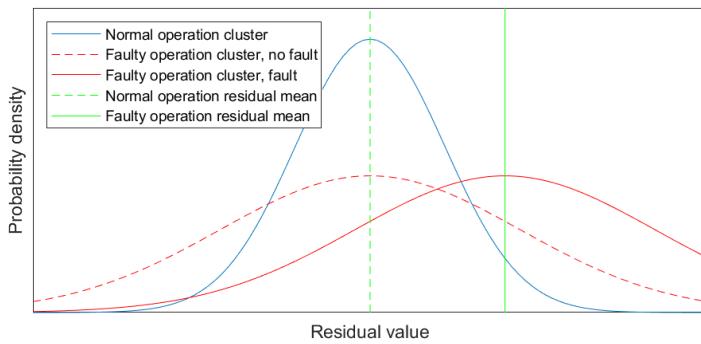


Figure 8.2: Illustration of fault tracking. The larger the mean value is, the more the faulty operation cluster tracks it

Finally, apart from the mixing coefficients, the only parameter being changed is the mean of the faulty operation cluster mean. If the variance of this cluster is larger, than the normal operation variance, its mean value has smaller density value. This means that, when the residual is more coloured, the faulty operation cluster takes more responsibility in the E step, hence it tracks

the residual more in the M step. When the residual has a mean, at which the normal operation cluster density is too low, the alarm is fired. This behaviour reduces the influence of the changes of variance on the fault tracking procedure. Therefore, if faulty data is available for training, more faults can be tracked with more clusters, leaving it possible to isolate the faults by tracking. However, in the investigated case, isolation is not considered, but the tracked faulty operation cluster mean can be used as a memory vector, which describes the correlation in the residual signal. Note from Figure 8.2, that in case of a fault, occasional deviations due to noise towards the other side of the normal operation curve - other side from the fault tracking cluster - still do not result in deterioration of the fault tracking, as the density values are only larger for the fault cluster, when this deviation is very large.

The iteration equations then become

$$\begin{aligned}\gamma(z_{n,k}) &:= \frac{\pi_k \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)}{\sum_{j=1}^K \pi_j \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)}, \\ \gamma(z_k)_N &:= \sum_{n=1}^N \gamma(z_{n,k}) \\ \boldsymbol{\mu}_k &:= \frac{1}{\gamma(z_k)_N} \sum_{n=1}^N (\gamma(z_{n,k}) \mathbf{x}_n) \\ \pi_k &:= \frac{\gamma(z_k)_N}{N}, \\ g &:= \pi_1,\end{aligned}\tag{8.12}$$

where N is the window length. The threshold becomes naturally

$$h_{\text{EM}} = 0.5,\tag{8.13}$$

as it indicates that the samples are more likely to be faulty.

The variance of the faulty operation Gaussian is found by tuning. The easiest closed form way found is setting the maximum density ratio at the mean of the normal operation cluster. This value is maximized, when the means are identical. It can be shown, that given the maximum density ratio MDR, the determinant of the variance of the faulty operation Gaussian cluster is given. Then, using the determinant rule $\det c\mathbf{A} = c^D \det \mathbf{A}$, the data having the dimension D, the MDR is set to the desired value with

$$\begin{aligned}\text{MDR} &= \frac{\mathcal{N}(\boldsymbol{\mu}_0 | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_1)}{\mathcal{N}(\boldsymbol{\mu}_0 | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)}, \\ |\boldsymbol{\Sigma}_1| &= \left(\frac{1}{\text{MDR}} - 1 \right)^2 |\boldsymbol{\Sigma}_0|, \\ \boldsymbol{\Sigma}_1 &\equiv \sqrt[D/2]{\frac{1}{\text{MDR}} - 1} \boldsymbol{\Sigma}_0.\end{aligned}\tag{8.14}$$

However, there are some pitfalls to correct during the implementation of the detector. Since the residual signal does not change its statistics for a long while, the responsibility of a cluster may go so low (due to the low window size), that the detection would hardly ever happen. Therefore the responsibilities were minimized at 0.001, and the density scores were minimized at 1E-20.

For this project, the Maximum Density Ratio is chosen as 0.1, since this provided the best trade off between false alarm and missed detection rates.

8.4 Fault handling

In this section, the fault handling is described. The main motivation is to detect the fault as soon as possible, then replace the measurements with estimations of the RLS algorithm. Therefore the operation falls into the area of fault accommodation [6]. At the same time, the algorithm shall be computationally efficient and not requiring too much memory.

The scenario that is pictured in case of a fault is the following. After the fault accommodation starts, it is likely to continue until a professional fixes the problem. Then, the professional can reset the fault handling algorithm. Therefore, even if the residuals settle around zero again, there is no need to detect switching back; however, estimations are still required to be exact, until the help arrives.

Figure 8.3 describes the chronology of the strategy. First, a relatively stable operation point is to be reached, to ease the optimization problem. Then the RLS algorithm is switched on. If batches are used, then the end of the first ambient temperature period is the starting point, since afterwards the variance tends to drop to a constant level, meaning that there is no more significant information in the model.

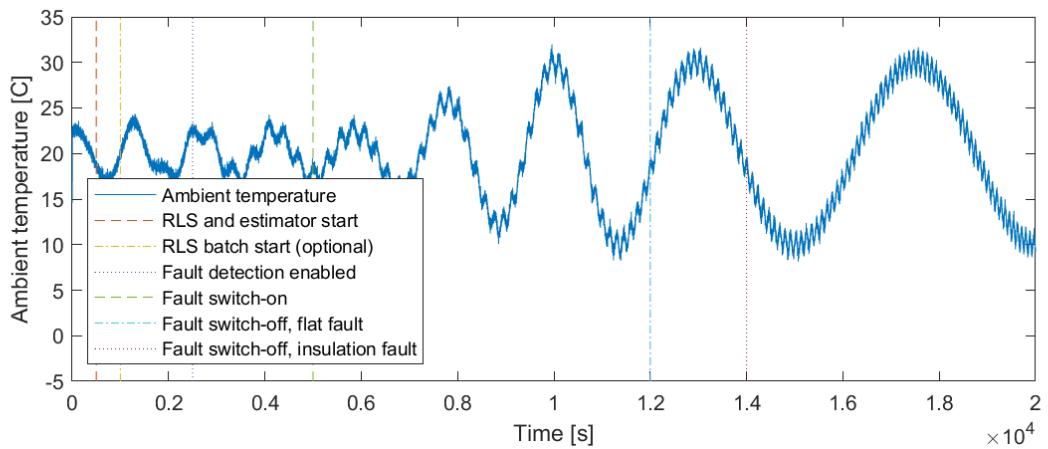


Figure 8.3: Fault handling events with the ambient temperature scenario (simulation)

When the moving average of 20 samples of decision function (of the EM algorithm) is greater than 0.9, the fault handling is switched on. The parameter vector 500 seconds before is saved, placed in a latch, and it is used from then to override the measurement (even if the fault ends). Due to the real life scenario described, fault handling is not switched off.

Chapter 9

Results and Evaluation

In this chapter, the fault handling closed loop system is compared to the 'ignorant' version (when the measurement has not been reconstructed). First, Section 9.1 shows evaluation statistics of the fault detectors, then the reconstruction of the measurements for the simulation. For the field data, since the original enthalpy is unknown, the fault is converted to temperature domain and it is evaluated there.

For the controllers, the fault estimation is used to modify the measurements, which are provided to the distributed controllers acting on the system. The metrics of the comparison is the coefficient of performance. For the estimators, the same values are used as estimator inputs. These are compared to the estimator results, when the fault is ignored. Both the fault handling and the 'ignorant' systems are investigated.

One of the motivations is to produce computationally cheap algorithms. This has been mostly achieved by using almost only linear operations, and the algorithms are only iterating along time-dimension. The largest computational challenges are carried out in the estimator, first, to take the exponential of a matrix in Section 6.1; second, to convert properties with the COOLPROP library. However, as it is described in Section 9.3, the iteration CPU time can be reduced by decreasing the set-point sampling time. Apart from the estimator, 20 exponentials are evaluated in the EM algorithm; nevertheless, the exponential function becomes computationally cheaper, when it considers a smaller number of the elements of the Taylor expansion.

9.1 Fault diagnosis and estimation evaluation

In this Section, the ways of evaluation statistics of the fault detectors are described, followed by the results. Then, the reconstruction of the measurements are depicted.

9.1.1 Evaluation statistics

The evaluation statistics of the fault detectors are

- Receiver operating characteristic curve (ROC), and area under the curve (AUC) value [48]. This indicates, how well a classifier separates the classes with its scores; in this case, how well the decision function values are separating the faulty from the normal operation cases. This evaluation is independent from the threshold.
- If the problem is perceived as the separation of two noisy signals along the threshold, namely the false positive rate and the true positive rate (will be described later), the sensitivity index [43] can describe this separation. For this statistics, the decision function is used again. This method is independent from the threshold.

-
- A less general, but more visual way of evaluation is investigating the missed detection and false alarm rates [6]. For this method, the fault evaluation signal is used. This method is dependent on the threshold, therefore reduces the domain of evaluation.
 - Fault switch-on or switch-off detection times are also dependent on the threshold, and measure how swiftly the fault evaluation signal is following the switches between the two operation modes.

The first two methods are giving more general ways of evaluating a classifier. However, the immediate requirements of the current application are, that the false alarm rates should be low, as well as the fault switch-on detection time. Note, that although these two qualities, and the general robustness (does the detector require retuning, when system parameters, or ambient temperature / fault scenarios are changed?) were taken into consideration, when the finally applied EM detector was chosen. The tuning of the detectors were based on the ideas described in Chapter 8, and were *not* dependent on the hereby presented evaluation metrics.

The ROC curve takes the true fault signal and the decision function, gradually moving a threshold from the minimum to the maximum value of the latter, while plotting the true positive rate (TPR) against the false positive rate (FPR). The TPR may be called as well: hit rate, detection rate, recall, or sensitivity. A synonym for FPR is false alarm rate. The expected ROC curve of a random classifier is a 45 degrees straight line. The AUC value is the area under the ROC curve, it ranges between 0 and 1, and the larger it is, the better the classifier is at the metric. One pitfall of it is that it gives badly scaled results for good detectors, with AUC values more than 0.95.

The sensitivity index or d' is a popular way of detector analysis in the field of signal processing. One advantage of it, is that it gives a better scaled evaluation for good detectors as well. It describes the mean separation between two noisy signals, weighted by the variances. For signal a and b ,

$$d' = \frac{\mu_a - \mu_b}{\sqrt{\frac{1}{2}(\sigma_a^2 + \sigma_b^2)}}, \quad (9.1)$$

which can be described by the rates of the ROC curve,

$$d' = Z(\text{TPR}) - Z(\text{FPR}) = \sqrt{2}Z(\text{AUC}), \quad (9.2)$$

where Z denotes the inverse of the cumulative distribution of the normal distribution function.

Diverging from normal distribution results in worse predictability of false alarm times or detection, or probabilities. Since the descriptive statistics, like Average Run Length (ARL) for the CUSUM algorithm, or probability of false alarm and missed detection for GLR are assuming both whiteness and normal distribution of the residual (see [6]), they are not considered.

9.1.2 Results

Figure 9.1 (fault handling on simulation data) and Figure 9.2 (no fault handling with field data) depict the residuals, either raw or whitened, the decision functions and the fault evaluations. The whitened residual is not corresponding to perfect white noise, and this is probably due to that linear inverse filters do not fit the nonlinear behaviours of signals. It turns out that the CUSUM algorithm is slowly reacting to switches, and that for the simulations, due to the low fault values (around 3 K), the EM algorithm has some missed detections, while for the field data, the GLR has some false alarms. This is more likely to be caused by the difficulty of the task, than by the performance of the detectors, since the rates can be changed by setting the thresholds. The detection time statistics are presented in Table 9.1, for the fault operational scenarios. Note that for the simulation with the insulation fault, the switch-off times are not represented, since due to the ramp (see Figure 5.3),

all the decision functions have already descended below the thresholds. Furthermore, the switch-off times are missing in the flat fault case as well, since the simulation crashed. At one occasion, a negative value is present in the table, since the detector was in the phase of false alarm for that amount of time.

	Simulation, flat fault	Simulation, insulation fault	Field data (on/off)
CUSUM	121 s	422 s	2 s / -
GLR	46 s	183 s	-120 s / 2 s
EM	39 s	383 s	16 s / 27 s
Benchmark	-	-	1072 s / 472 s

Table 9.1: Detection times of detectors. Negative value indicates current false alarm length

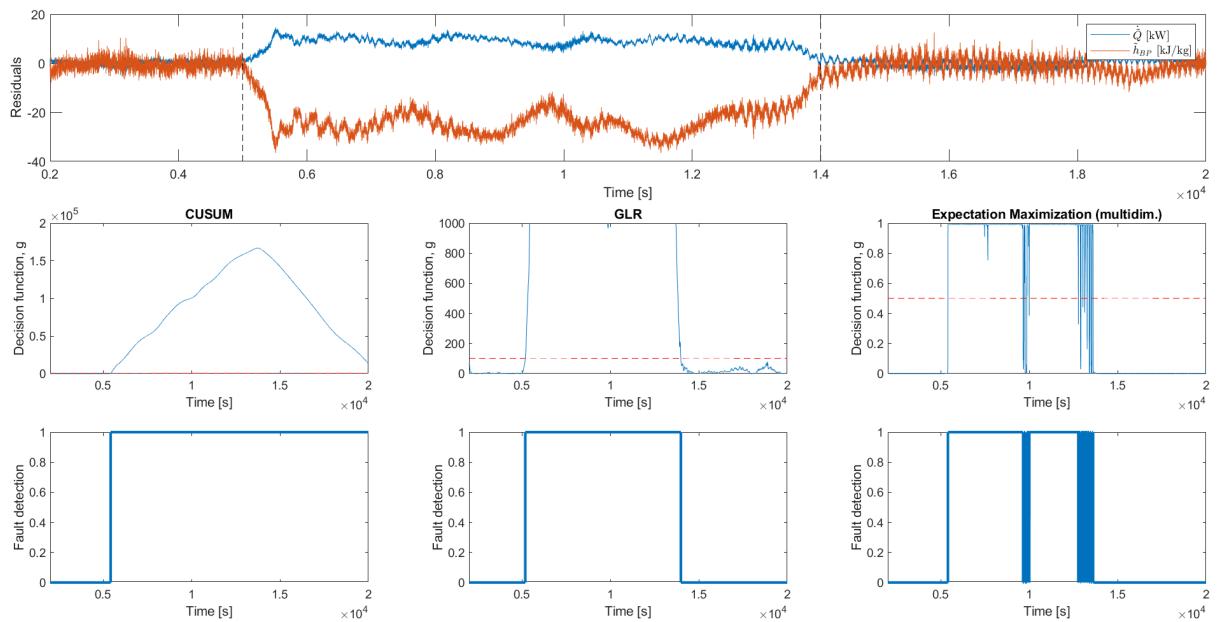


Figure 9.1: Residuals and fault detection for simulation, in case of insulation fault

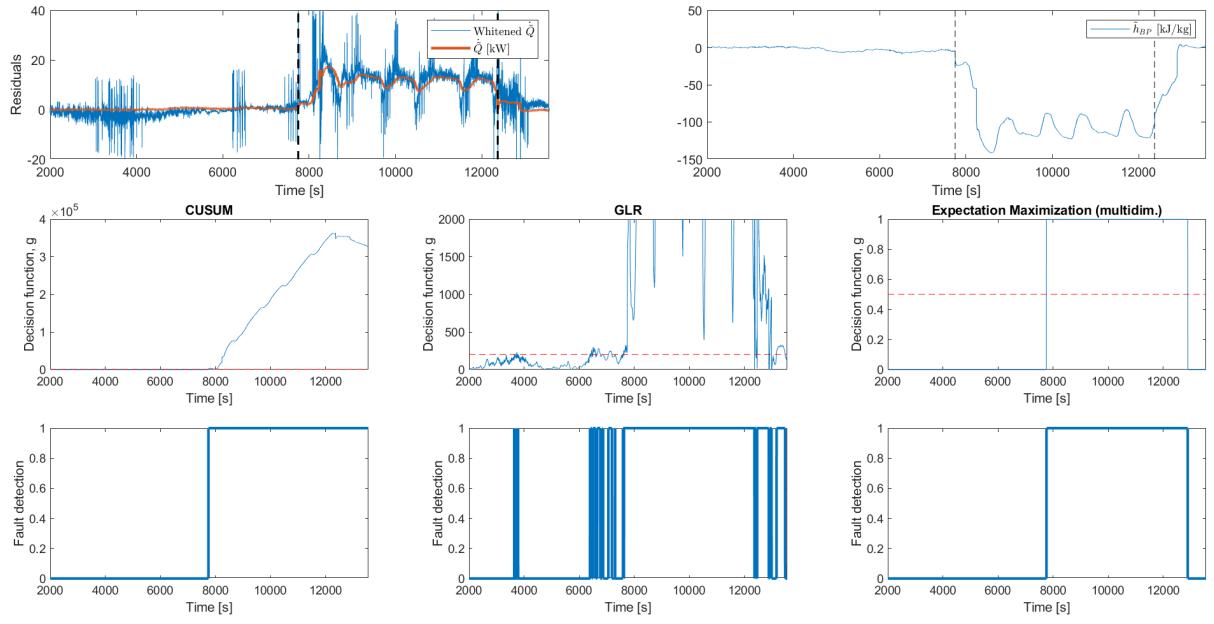


Figure 9.2: Residuals and fault detection for field data

Table 9.2 presents the missed detections, and Table 9.3 the false alarms. Only the insulation fault scenario is considered for the simulation. This is due to the fact, that the data set there is twice as large, therefore it gives better statistics. What becomes clear immediately, is that for the given residual, the CUSUM algorithm is giving a lot of false alarms. This is due to its slowness, see Figure 9.1. Furthermore, the designed GLR seems to be more sensitive to mean deviations, than the designed EM, but as it has been mentioned, this is based on the choice of threshold.

	Simulation, fault ignored	Simulation, Fault handling	Field data
CUSUM	4.7%	4.7%	0.0%
GLR	4.8%	2.2%	0%
EM	4.4%	12.1%	0.3%

Table 9.2: Missed detection rates, in case of handled insulation fault

	Simulation, fault ignored	Simulation, Fault handling	Field data
CUSUM	60.0%	60.0%	15.0%
GLR	10.6%	14.6%	18.6%
EM	12.0%	4.5%	6.6%

Table 9.3: False alarm rates, in case of handled insulation fault

Figure 9.3 visualizes the ROC curves and their AUC values for the detector decision functions, and the residuals. Table 9.4 shows the sensitivity index values calculated by Eq. 9.2. The residual of \hat{h}_{BP} is the best classifier. This is due to the fact, that the signal-to-noise ratio (SNR) [20] is large for the signal. This further indicates, that the CUSUM and the GLR detectors might have got the worse residual as input, at least for static classifications; however, the response time of this signal was too large for fast detection times. The GLR algorithm proves to be a slightly

better static classifier, than the EM. Another finding, is that the field data residual is generally a better residual, since the fault was a step signal there, giving better separation of the mean values.

Note that there are important limitations of the evaluation statistics, which are

- The residuals, hence the decision function are dependent on the fault and the ambient temperature scenarios.
- In case of field experiment data, the CUSUM and the GLR use the whitened versions of only the residual of \hat{Q} . This whitening, as picked with naked eye, did not ease the detection task (see Figure 9.2). Furthermore, the EM algorithm uses two signals for detection, and the additional one became the best classifier according to threshold-free statistics. However, more signals do not only introduce more information, but the curse of dimensionality as well. What can surely be stated, is that a more equal comparison should be done.
- One of the most important statistics, the detection time, is not taken into consideration in the AUC value, since that metric is static. Hence one single evaluation method is not eligible to evaluate the detectors.
- The window length of the GLR and EM detectors were different; it was 100 for the GLR algorithm (this was the only way to make it work), and 10 for the EM algorithm.

In conclusion, there is one clear result of this section: the CUSUM algorithm is not sufficiently fit for the presented fault diagnosis problem.

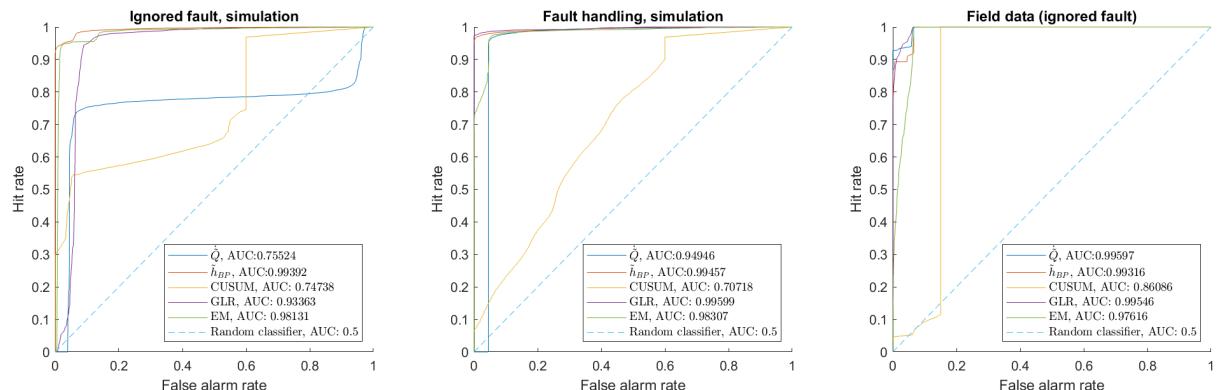


Figure 9.3: Receiver operating characteristic curves. For simulation evaluation, insulation fault is considered

	Simulation, fault ignored	Simulation, Fault handling	Field data
\hat{Q}	0.97	2.32	3.74
\hat{h}_{BP}	3.55	3.60	3.48
CUSUM	0.94	0.77	1.53
GLR	2.13	3.75	3.69
EM	2.94	3.00	2.80

Table 9.4: Sensitivity indices, in case of handled insulation fault

9.1.3 Reconstruction of measurement

The reconstructions of the faulty measurement signal for simulation are presented in Figure 9.4 for the case of the ignored fault, and in Figure 9.5 for the fault handling case. It can be seen

that the reconstruction results are much better, if the fault is detected in time and the system stays in normal operation domain. Worse reconstruction for the former scenario can be caused by either saturation of a signal, or discovering new operation points, or both. The residual \hat{h}_{BP} is overestimating, while $h_{HR} - \hat{Q}/\dot{m}_{HR}$ is underestimating the fault. Since the former gives better reflection of the true dynamics, and it turned out to be more robust during further experiments, it is chosen the signal to override the measurement. Due to feedback of the noisy reconstructions, the variance of the measurement signal is larger after the switch-off, than during the first normal operation scenario.

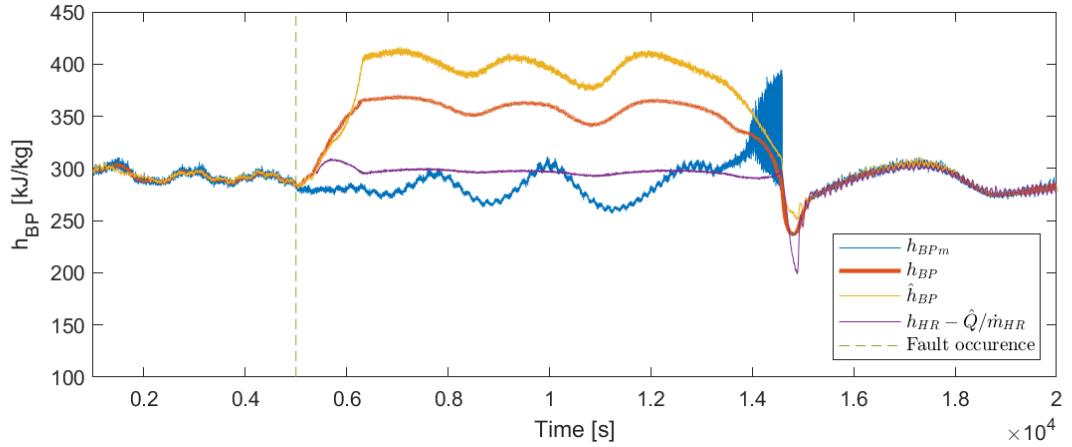


Figure 9.4: Reconstruction of outputs for simulation, when insulation fault is ignored

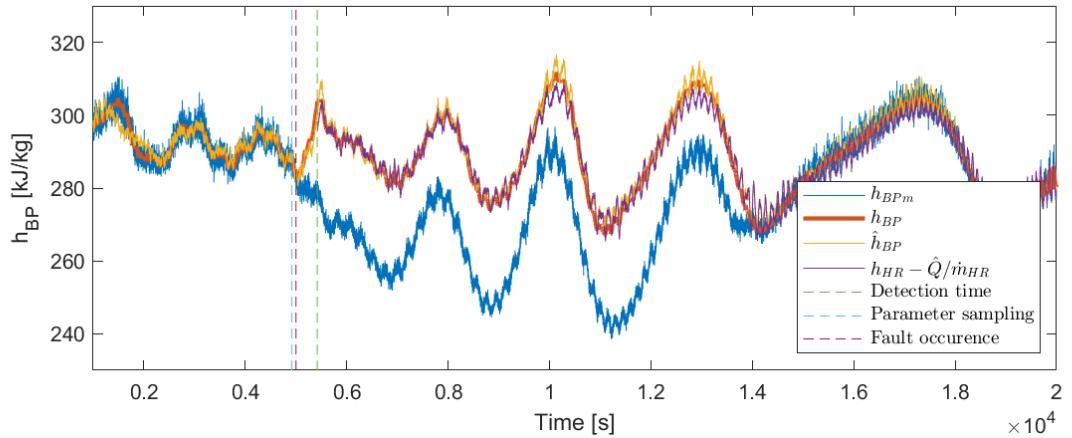


Figure 9.5: Reconstruction of outputs for simulation, with insulation fault and fault handling

The enthalpy measurement reconstruction is used to calculate faults in temperature domain in Figure 9.6. Since the more resolute (see the ph-diagram in Figure 3.1) enthalpy domain provided good results, the temperature domain will provide good results as well.

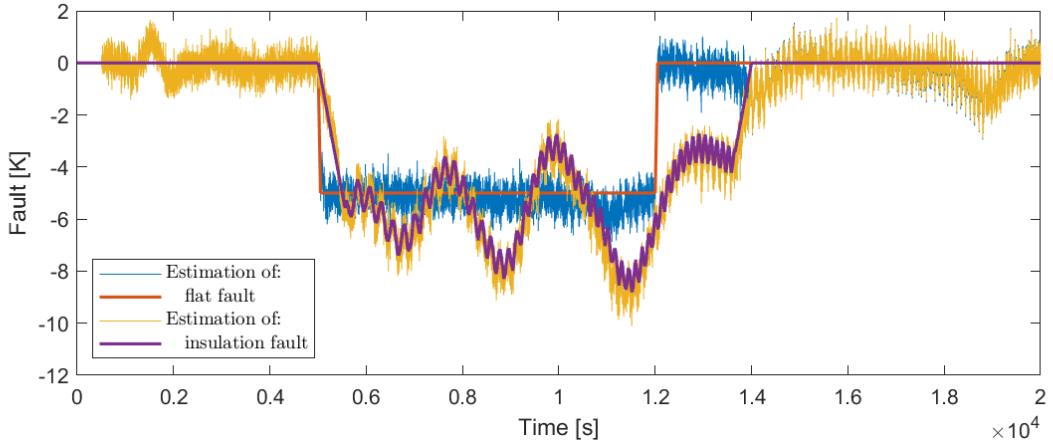


Figure 9.6: Fault estimation for simulation. Fault handling is active

The reconstruction plot cannot be depicted for the field data simulation, as sometimes the enthalpy falls under the saturation curve, hence the true value cannot be calculated, even if the fault is known. However, converting the results to temperature domain, comparison can be done again. Figure 9.7 presents the results. After small transients, the fault is estimated with about 30 % estimation error, then some bias can be observed during the reset normal operation. Note however, that the latter may be caused by transient effects (which can be clearly seen on the input signals). Generally, the training information content was much less, than the one for the simulation, hence worse results are reasonable. In conclusion, for better estimates, the RLS algorithm should be trained through a sufficient scale of data. On the other hand, note that training on 25 % of the test amplitude of the ambient temperature (see Figure 5.2) provides decent results for the simulation data.

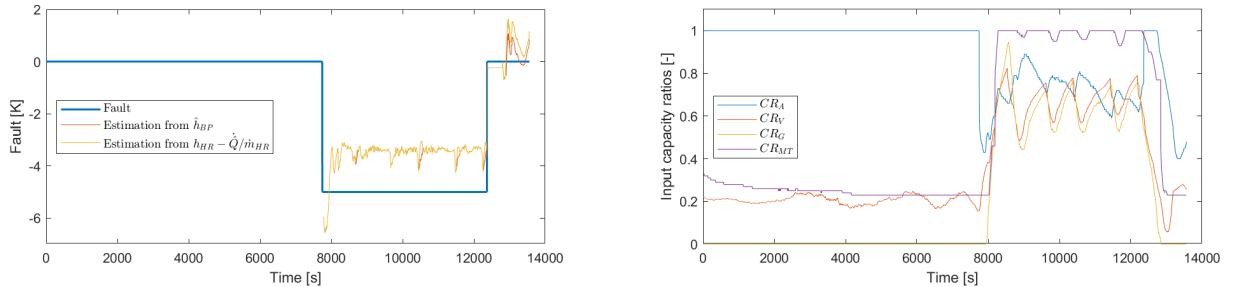


Figure 9.7: Fault and inputs for field data

One finding is that the mean of the residual, hence the detectability is dependent on the linearization point. It is easier to detect in case of low ambient temperature. The enthalpy at the outlet of the gas cooler is lower, where, as seen in the ph-diagram (Figure 3.1), the partial differential $\frac{\partial h}{\partial T}$ is large. This has direct effect on both of the residuals. This is the same area, where fault detection is critical, since the effect is the harshest at this region, for the same reason.

9.2 Control performance analysis

In this section, the performance of the fault-tolerant controller is compared to the one, which does not consider the fault. The metric of the comparison is the Coefficient of Performance.

9.2.1 Coefficient of performance

The coefficient of performance for a cooling system is defined as the ratio of two energy quantities; the energy transported from the evaporators, divided by the work, which is used actively to do that (see [19]). For time-varying systems, it can be evaluated as a time-varying function, with the derivatives of these quantities. Hereby it is defined as

$$\text{COP} = \frac{|\dot{Q}_{\text{evap}}|}{\dot{W}_t} = \frac{\dot{m}_C(h_C - h_L) + \dot{m}_F(h_F - h_L)}{\dot{m}_{\text{MT}}(h_{\text{MT}} - h_{\text{MT,in}}) + \dot{m}_{\text{MT}}(h_{\text{MT}} - h_{\text{MT,in}}) + \dot{m}_{\text{MT}}(h_{\text{MT}} - h_{\text{MT,in}}) + \rho_A \dot{V}_A(h_A - h_{\text{A,in}})}, \quad (9.3)$$

where \dot{Q}_{evap} is the heat flow rate of the evaporator, index F denotes freezer outlet, index C cooler outlet, index MT medium temperature and index LT low temperature compressor, index IT parallel compressor, index in inlet. The enthalpies of the compressors are understood as their outlet.

9.2.2 Results

The original fault handling control strategy is presented in Figure 9.7. After the IT compressors or the receiver valves are saturated for a while (since the receiver pressure increases), the pressure is raised to increase heat transfer. The saturation region is left, the fault handling is switched off. However, the faulty sensor measurement does not vanish, and brings the system into fault handling again. This oscillatory behaviour does not result in as good COP values as earlier, since the compressors need to transfer refrigerant to a higher pressure level.

The two main effects of gas loop are dropping COP and challenging control of the evaporator outlet superheat. The control of the latter is presented in enthalpy domain, since at the given pressure level the saturation curve is almost vertical and the isothermal curves are parallel.

Figure 9.8 depicts that for the same ambient temperature scenarios, the COP level of the fault handling system should keep the normal pattern in case of the fault occurrence at 5000 seconds, while the fault ignorance should result in decreasing values. This effect is present in the figure. The flat fault simulation did not finish, since the nonlinear differential equations of the mathematical model became unstable. Therefore, in the forthcoming results, the insulation fault scenario will be presented, which is also more realistic. The fault handling curves are identical after some small transients of 500 seconds, which indicates that neither of the faults have any effect on the control performance. After the switch off time at 14000 seconds, the signals become identical in less than 2000 seconds. Note that the COP is always higher at lower ambient temperatures, due to the increasing heat flow rate, which results from the higher temperature difference for the same pressure level.

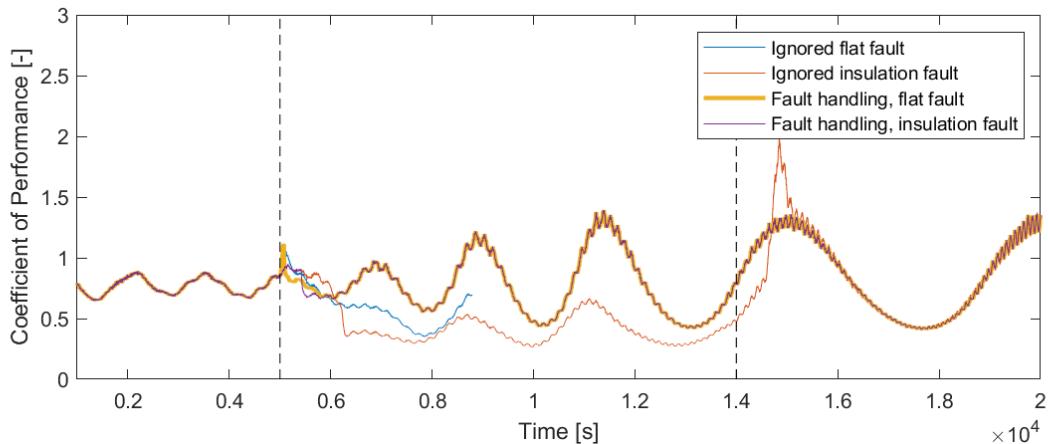


Figure 9.8: Coefficient of Performance (COP). Fault switch-on and switch-off are indicated by vertical lines. In case of ignored flat fault, the simulation crashed at ≈ 9000 s

Note that the COP values are generally low for the simulated plant. In the industry, it ranges between 1 and 4 for carbon-dioxide systems [15]. This may be due to two reasons. The first is that the work done by the fans is considered in Eq. 9.3, which is not done in the original definition [11]. The second is that the industrial controllers include a rule-base apart from the distributed PI controllers, mostly to avoid control synchronization. The synchronization problem is described in [36].

In Figure 9.9, the evaporator outlet enthalpies challenge the fault ignoring controllers after the fault occurrence. This value is crucial to be constant, so that the cabinet temperature can be robustly controlled. There is a delay of about 500 seconds after the fault for the oscillations to pop up, and the decaying lasts more than 2000 seconds after the switch-off time. The reason, why the system is not exploding, is the natural stability of thermodynamical systems: the energy accumulation results in larger heat-flow rates, even, if the dynamics are uncontrollable.

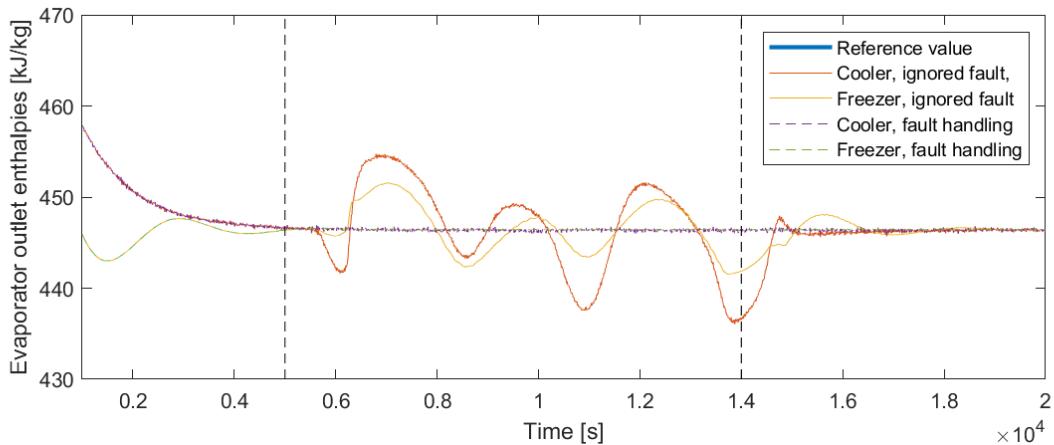


Figure 9.9: Evaporator outlet enthalpies. Fault switch-on and switch off are indicated with vertical lines

In Figure 9.10, the input signals are depicted for the fault ignoring and the fault handling cases. It turns out that in case of fault handling, not only saturation, but any unusual controller behaviour (for example spikes) is avoided.

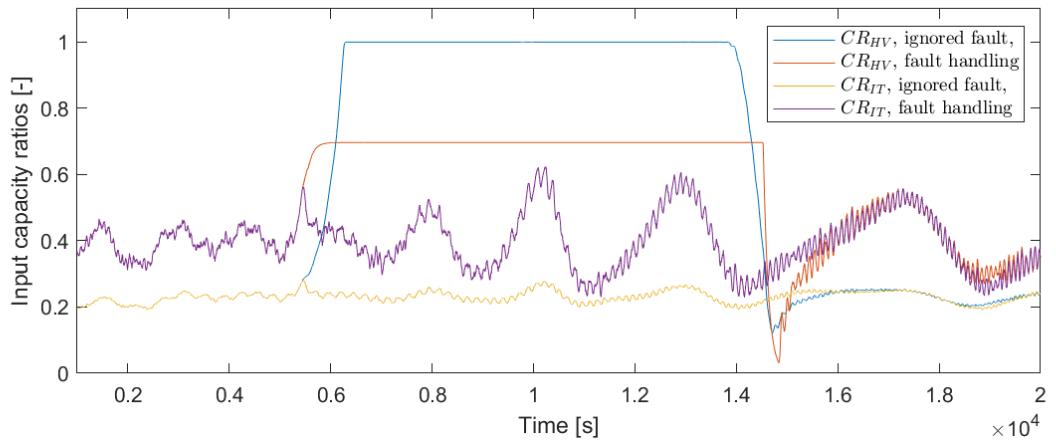


Figure 9.10: Input signals, simulation data, insulation fault

9.3 State estimator evaluation

In Chapter 6, it is shown that the innovation signal of the linear estimator cannot be used for fault diagnosis. In Chapter 7, a method for estimating the fault is described, for which the fault handling strategy is presented in Section 8.4. If the estimator is to be used for estimation of other physical properties, or predictive control [29], it shall use these results. Furthermore, a working estimator is one more applicability test for the fault estimation.

Most of the states of the system are measured. Due to the large Kalman gain chosen to track the states and hence the set-points quickly, the estimation errors of these states are mainly the function of the measurement noises. Density values are set by the set-point selector as described in Section 6.6. Therefore only the virtual states are presented, in Figure 9.11 and Figure 9.12. The former shows the virtual gas cooler enthalpy, the latter the virtual air temperature in the gas cooler. Note that either in case of fault handling or fault ignorant control strategy, an estimator considering, and an estimator ignoring the fault are presented.

The enthalpy states are well estimated with and without fault handling. It seems, that the errors found in Figure 9.4 are filtered out. Note that although the program running time improves in case of less frequent set-point update (which includes inversion of large matrices), the results are less correct. For numerical results of running time, see Table 9.5.

Set-point sampling time	Sim., ignored fault	Sim., fault handling	Field data
100 s	4.4 ms	4.3 ms	3.5 ms
1000 s	3.5 ms	3.2 ms	2.6 ms

Table 9.5: CPU running time of 1 second simulated estimation time
(fault handling is applied)

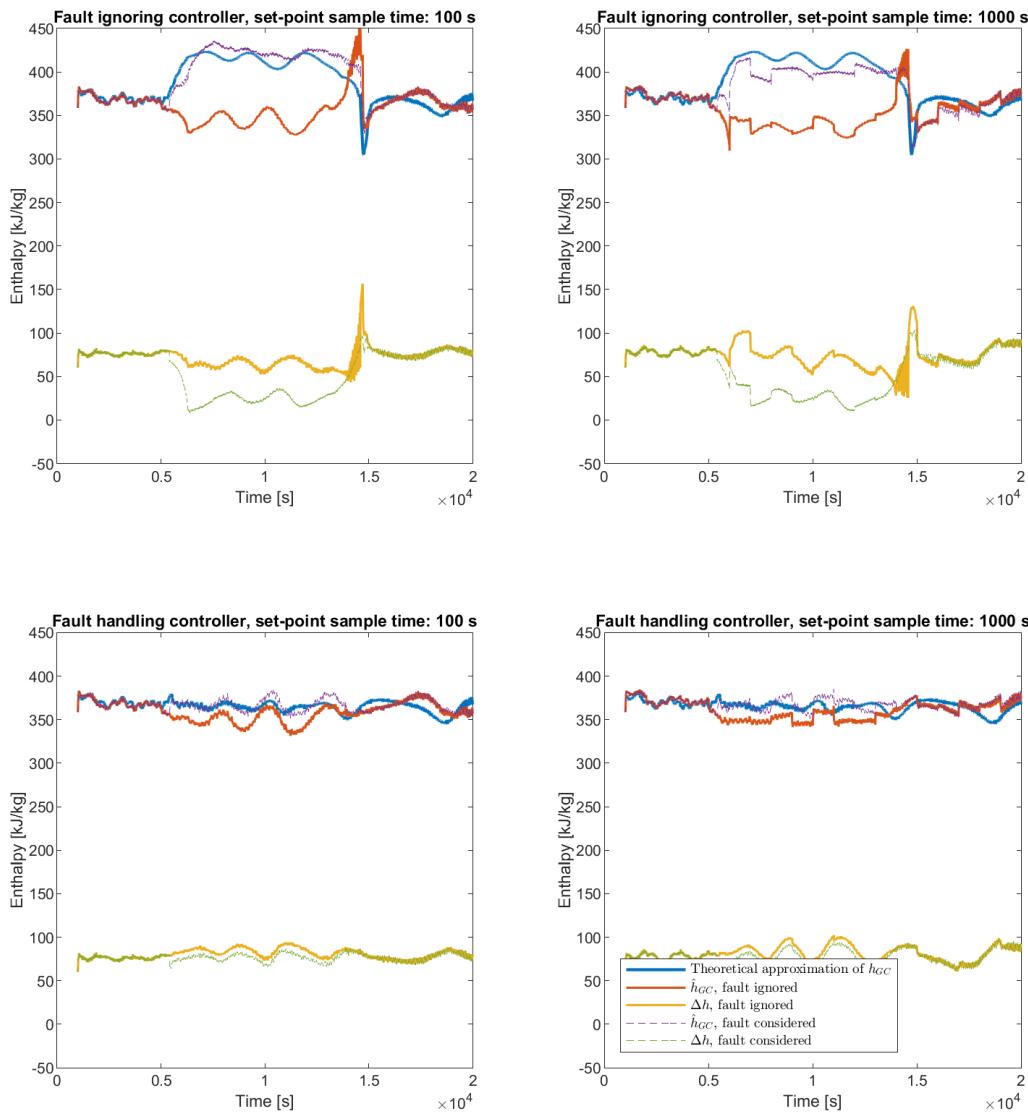


Figure 9.11: Estimation results for virtual gas cooler enthalpy, in case of insulation fault. The results of the fault accommodating estimator are denoted by dashed lines

The results in Figure 9.12 are less straightforward to evaluate. Note the difference of the scaling. The analytical calculation of the virtual air temperature state is

$$T_{A,GC} \approx T_{GC}(p_{GC}, h_{GC})_{ss} - 3[K], \quad (9.4)$$

which is very inaccurate, as the temperature difference is not a single value. It ranges between 2 and 4 degrees in time at the outlet of the refrigerant side of the gas cooler, and along the gas cooler tube, as it was depicted by simulation results of the special course (see Figure 3.7). The estimator gives a much better estimate, if the fault is known. However, even knowing about the inaccuracy, the errors are large. In Figure 9.13, the estimations are compared to the analytical calculations. It can be seen that the more frequent the set point change is, the smaller the estimation error is. On the other hand, some phase lead is indicated in the diagrams, since the area enclosed by the curves are larger. This can be most probably reasoned by Eq. 3.14. Either the time constant underestimated, or the value of the resistance ratio w is overestimated, therefore the model 'predicts' the state.

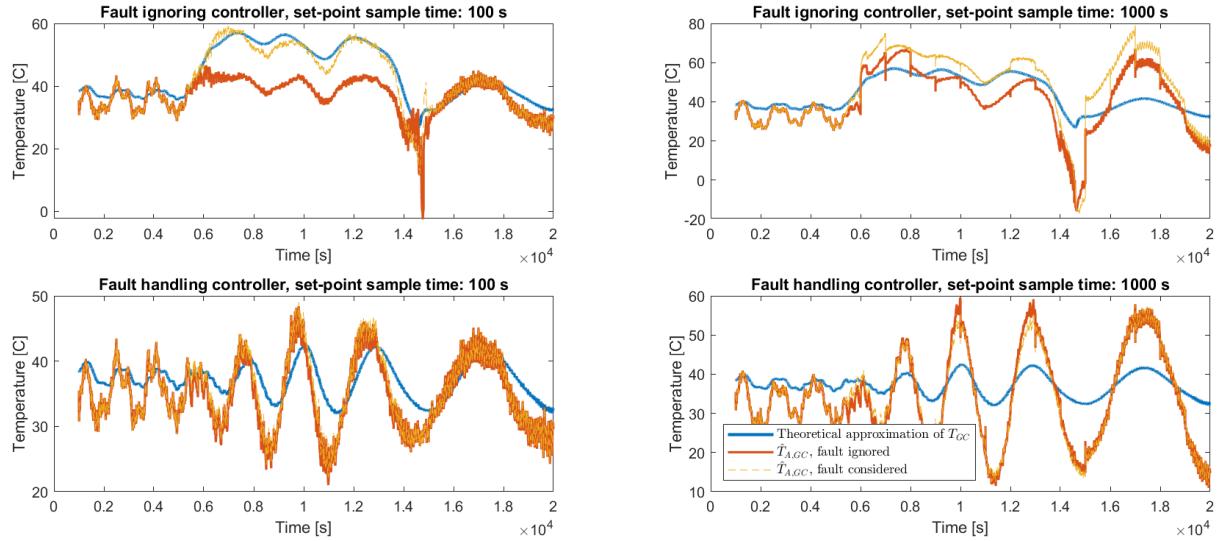


Figure 9.12: Estimation results for virtual gas cooler air temperature, in case of insulation fault. The results of the fault accommodating estimator are denoted by dashed line

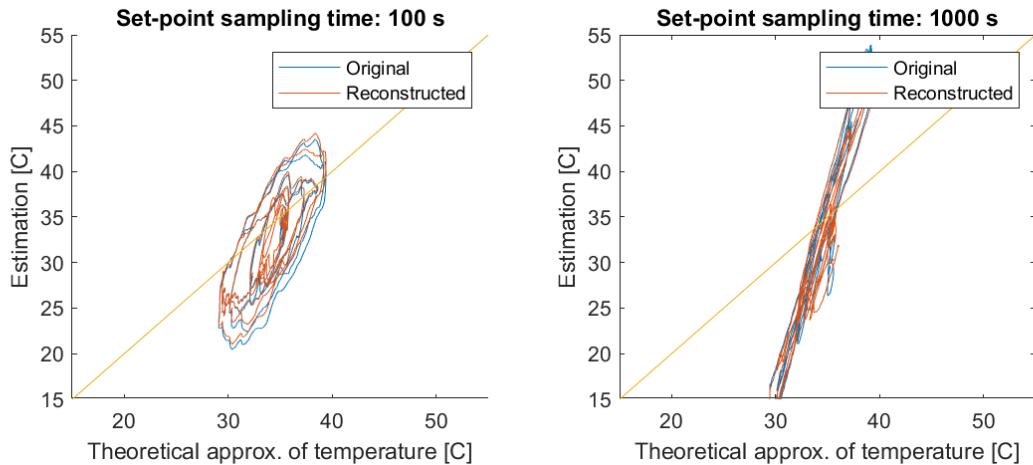


Figure 9.13: Estimation results against theoretical approximations, in case of insulation fault and fault handling. Periodically enclosed area indicates phase lead or lag

In Figure 9.14, the unfortunate pattern is shown, which occurs, while using the constrainer described in Section 6.5. When the fault is ignored, the enthalpy inlet of the gas cooler is perceived as a low value. Therefore, the constrainer sets the value of the receiver enthalpy low at every set point selection, as it should. However, the measurements of the receiver filling level are pulling the values back through the large Kalman gain. In conclusion, both the constrainer and the estimator are working adequately, the former avoiding unstable effects or CoolProp library crashing due to physically impossible property-combinations, the latter minimizing the estimation error variance.

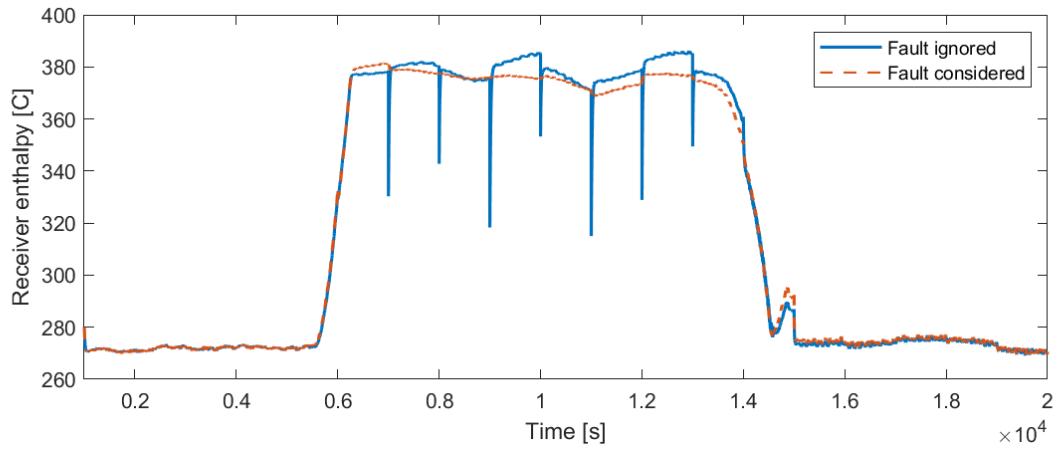


Figure 9.14: Results of constraints for the enthalpy of the receiver, in case of insulation fault and fault ignoring controller

The state estimation results for the field data cannot be validated, since there is no additional information to analytically estimate the virtual states, how it is done for the simulation data. Therefore they are presented in Figure F.11 in Appendix F. One achievement of it, is that the constrainer managed to keep the receiver enthalpy lower, than its inlet enthalpy.

Chapter 10

Conclusion and Discussion

10.1 Thesis summary

The proposed objective of this thesis was to solve a common industrial problem of refrigeration systems, namely gas loop, which is caused by faulty sensor measurement, and has negative effects on the entire system, like decreasing efficiency, or products heating up. The requirement is to find a strategy to prevent these .

The chosen strategy was to do early fault detection, then with measurement reconstruction, prevent gas loop. In order to provide an applicable strategy, robustness, computational price, and modularity are considered as well. For the sake of modularity, the fault diagnosis problem was divided into residual generation and fault detection.

The main challenge was creating a robust strategy. The signals are dynamical, the measurements are corrupted by non-Gaussian noise, the model carries uncertainties, and the system has virtually infinitely many operation points along a given operation domain. It turns out that the fault diagnosis problem is eased, if the ambient temperature is lower, on the other hand, in case of missed detection, the effects are more pronounced.

Finally, the residuals are generated by the Recursive Least Squares algorithm, assuming linear relation between the energy flows. The energy accumulation is bounded and this bound can be robustly captured, when exploring the operation domain. For fault detection, the conventional online algorithms are replaced by a modification of the Expectation Maximization algorithm. The algorithms are computationally cheap. The fault diagnosis method uses only measurements and input signals, therefore modularity was guaranteed. A complex simulation scenario was used to show robustness. Closed loop simulations, and an open loop field experiment were used to demonstrate, that the detection of the fault was early enough to avoid gas loop.

10.2 Conclusions

The first contribution of this thesis pertained to the derivation of a nonlinear model, based on physical first principles. This model was simplified, until only the most important traits of the system were captured. The model was linearized, and tested at a selected operation point, where it was shown, that the linearized model reflects the behaviours of the nonlinear one, and that it can be used for control and estimation purposes. Look-up tables were generated for linearizations along the operation domain.

The resulting model was tested against two realized systems: a high fidelity simulation programmed in Modelica language, and a field experimental setup at the company Danfoss. The

former was used, so that the gas loop phenomenom can be captured, and could be manipulated to consider the faults. The latter was an experiment, where the fault was not handled during the operation. For the simulation data, the fault and ambient temperature scenarios, and their motivations of testing robustness were presented.

First, a two-fold discrete state estimation filter was used to produce residual signals, as its innovation. The outer fold was setting the stationary states, considering steady-state property conversions, and took account of extreme values. The filtered values were used together with look-up tables, to deduce the parameter matrices for the set-points. In the inner fold, a linear Kalman filter was used, to estimate local dynamics. It was shown, that the nonlinearities have the same effect on the residuals as the fault. This is due to the fact, that the large Kalman gain, which was required for tracking the set-point well, weakened the detectability. Therefore a new method was to be found.

The online and computationally cheap Recursive Least Squares algorithm was proposed, assuming that the energy accumulation of the gas cooler was bounded. The method was independent of the state estimator, for stability and modularity reasons. Regularization methods were used, until the algorithm became robust enough to perform well on the simulation scenarios. Strong detectability was achieved by freezing the parameter vector after the fault detection. For the field data, residual whitening was designed based on information criteria, F-test, and further methods, and provided whiter, but less normally distributed residual.

The fault detection was done by three algorithms. The CUSUM and GLR algorithms were using a single-dimensional whitened residual signal. The third algorithm was created to be able to track and possibly isolate the fault, and robustly use the multidimensional coloured residual signal. For this purpose, the Expectation Maximization algorithm for Gaussian Mixture Models was converted to its online and simplified version. For all the three algorithms, the fault detection was around one minute, which is less than the the benchmark of fifteen minutes for the same scenario. A robust fault handling strategy was designed, to keep the plant working, until human intervention.

The results were presented, including both fault ignoring and handling scenarios. The latter used the Expectation Maximization algorithm. The simulation scenarios were designed as stress-tests well representing ambient temperature time evolution, while being simulation-efficient. The three areas of evaluation were measurement reconstruction, showing that the estimation of the RLS algorithm was keeping close to the true values; control performance analysis, where it was shown that the fault has no effect on the COP value or the cabinet temperatures in case of fault handling; and the state estimator, showing that there was only one state, for which the estimation may make compromises with the applications. Furthermore, the fault detectors were compared based on their ROC curves, false alarm rates, detection times, and further analyses. The presented methods handled well the information-poor scenario of the field data experiment.

10.3 Discussion and future work

As a next step, the presented fault detection, measurement reconstruction and fault handling strategies could be tested within field data experiments with fault handling. This may bring up robustness issues, which are not accounted for. Furthermore, more field data experiment would validate the applicability for information-poor scenarios for the Recursive Least Squares algorithm. If the fault estimators designed are not sufficient, the online fault profile estimation method of [44] may be combined with this one, letting the knowledge of the fault to be further evolving during fault handling as well.

In order to make deeper analysis, and since only this could be validated through field exper-

iment, the scope of the project was kept from 10 to 30 °C. A further continuation of the project could be extending the found strategies to covering lower ambient temperatures, when the bypass filter and heat recovery units are used. As it was found, detectability was expected to be higher for the RLS algorithm for this range. On the other hand, at this range, observability problems may occur, when the gas cooler outlet enthalpy enters the region under the saturation curve, for which situation the estimator was not prepared.

In this work, a multidimensional online fault detection algorithm was designed, based on the EM algorithm; however, due to the complexity of the task of the detection challenge, well-based comparison could not be set up to test its performance. An evaluation should be done on it, with less complicated test scenarios, when it receives the same residual signal as the other detectors. The bases of evaluation could be: detection time, alarm rates, AUC value, sensitivity index, memory usage, CPU running time, and possibly robustness to environment changes.

The designed estimator was used for validation. However, there is more potential in it. It provides enthalpy values, which are much better resolved above the saturation curve of the pressure-enthalpy diagram, and provide more information of the energy content of the refrigerant under it. This, being used for control purposes, may result in better COP values.

The state estimator could be used to enhance modern control strategies. The current strategy in industry is distributed control with a rule-base. In case of altering the plants, this method is challenged, and the applied proportional-integrator (PI) controllers are to be re-tuned. However, the designed state space description includes the most important traits of the back-end of the refrigeration system, and if a Kalman predictor is used, the model itself could be used for predictive control. Since having a longer horizon allows more advanced control objectives and soft constraints to be applied, there is a potential for improving efficiency again.

Bibliography

- [1] Ken Aho, DeWayne Derryberry, and Teri Peterson. “Model selection for ecologists: the world-views of AIC and BIC”. In: *Ecology* 95.3 (2014), pp. 631–636.
- [2] Mariela De Lucas Alvarez and David M Lane. “A Hidden Markov Model application with Gaussian Mixture emissions for fault detection and diagnosis on a simulated AUV platform”. In: *OCEANS 2016 MTS/IEEE Monterey*. IEEE. 2016, pp. 1–4.
- [3] RA Belcher. “Multi-tone testing of quantisers using PRBS signals”. In: *Computer standards & interfaces* 22.4 (2000), pp. 269–279.
- [4] Ian H Bell et al. “Pure and pseudo-pure fluid thermophysical property evaluation and the open-source thermophysical property library CoolProp”. In: *Industrial & engineering chemistry research* 53.6 (2014), pp. 2498–2508.
- [5] Christopher M Bishop. *Pattern recognition and machine learning*. Springer, 2006.
- [6] Mogens Blanke et al. *Diagnosis and fault-tolerant control*. Vol. 2. Springer, 2006.
- [7] Torsten Blochwitz et al. “The functional mockup interface for tool independent exchange of simulation models”. In: *Proceedings of the 8th International Modelica Conference; March 20th-22nd; Technical Univeristy; Dresden; Germany*. 063. Linköping University Electronic Press. 2011, pp. 105–114.
- [8] Fredrik Bagge Carlson, Anders Robertsson, and Rolf Johansson. “Identification of LTV dynamical models with smooth or discontinuous time evolution by means of convex optimization”. In: *2018 IEEE 14th International Conference on Control and Automation (ICCA)*. IEEE. 2018, pp. 654–661.
- [9] Bruno YK De Carvalho, Cláudio Melo, and Roberto H Pereira. “A Study on the Use of Liquid Separators in Variable-Speed Small-Scale Carbon Dioxide Refrigerating Sytems”. In: *12th IIR Gustav Lorentzen Natural Working Fluids Conference, Edinburgh–United Kingdom*. 2016.
- [10] Alberto Cavallini et al. “Two-stage transcritical carbon dioxide cycle optimisation: A theoretical and experimental analysis”. In: *International Journal of Refrigeration* 28.8 (2005), pp. 1274–1283.
- [11] Yunus A Cengel and Michael A Boles. “Thermodynamics: an engineering approach”. In: *Sea* 1000 (2002), p. 8862.
- [12] Lin Chen, Bi-Li Deng, and Xin-Rong Zhang. “Experimental study of trans-critical and supercritical CO₂ natural circulation flow in a closed loop”. In: *Applied Thermal Engineering* 59.1-2 (2013), pp. 1–13.
- [13] J Cioffi and Thomas Kailath. “Fast, recursive-least-squares transversal filters for adaptive filtering”. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 32.2 (1984), pp. 304–337.
- [14] Scott D Cohen, Alan C Hindmarsh, and Paul F Dubois. “CVODE, a stiff/nonstiff ODE solver in C”. In: *Computers in physics* 10.2 (1996), pp. 138–143.
- [15] “Confidential discussions with industrial partner”. In: 2019.

-
- [16] Abderrahim Frih, Zakaria Chah, and Mostafa Mrabti. “Controllability and observability of LTV systems-bond graph approach”. In: *Asian-European Journal of Mathematics* 11.03 (2018), p. 1850038.
 - [17] YT Ge and SA Tassou. “Thermodynamic analysis of transcritical CO₂ booster refrigeration systems in supermarket”. In: *Energy Conversion and Management* 52.4 (2011), pp. 1868–1875.
 - [18] Eric Granryd. *Refrigerating engineering*. Royal Institute of Technology-Department of Energy Technology, 2005.
 - [19] Torben Green, Roozbeh Izadi-Zamanabadi, and Henrik Niemann. “On the choice of performance assessment criteria and their impact on the overall system performance—The refrigeration system case study”. In: *2010 Conference on Control and Fault-Tolerant Systems (SysTol)*. IEEE. 2010, pp. 624–629.
 - [20] Md Ekramul Hamid et al. “An Improved Sad Algorithm of Noisy Speech Using Estimated Degree of Noise”. In: *canadian journal on signal processing issn* 1709 (1923).
 - [21] Elbert Hendricks, Ole Jannerup, and Paul Haase Sørensen. *Linear systems control: deterministic and stochastic methods*. Springer, 2008.
 - [22] Bowen Hou et al. “A new fault diagnosis method based on component-wise expectation-maximization algorithm and k-means algorithm”. In: *2017 6th Data Driven Control and Learning Systems (DDCLS)*. IEEE. 2017, pp. 776–781.
 - [23] Tobias Gybel Hovgaard et al. “Active sensor configuration validation for refrigeration systems”. In: *Proceedings of the 2010 American Control Conference*. IEEE. 2010, pp. 3604–3610.
 - [24] Yunpeng Hu et al. “Chiller sensor fault detection using a self-adaptive principal component analysis method”. In: *Energy and buildings* 54 (2012), pp. 252–258.
 - [25] Yunpeng Hu et al. “Sensitivity analysis for PCA-based chiller sensor fault detection”. In: *International journal of refrigeration* 63 (2016), pp. 133–143.
 - [26] Rolf Isermann. *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Springer Science & Business Media, 2006.
 - [27] Richard M Johnstone et al. “Exponential convergence of recursive least squares with exponential forgetting factor”. In: *Systems & Control Letters* 2.2 (1982), pp. 77–82.
 - [28] John Bagterp Jørgensen. “Adjoint sensitivity results for predictive control, state-and parameter-estimation with nonlinear models”. In: *2007 European Control Conference (ECC)*. IEEE. 2007, pp. 3649–3656.
 - [29] Jer-Nan Juang et al. “Identification of observer/Kalman filter Markov parameters-Theory and experiments”. In: *Journal of Guidance, Control, and Dynamics* 16.2 (1993), pp. 320–329.
 - [30] Minsung Kim et al. “Design of a steady-state detector for fault detection and diagnosis of a residential air conditioner”. In: *International journal of refrigeration* 31.5 (2008), pp. 790–799.
 - [31] Woohyun Kim and James E Braun. “Development and evaluation of virtual refrigerant mass flow sensors for fault detection and diagnostics”. In: *International Journal of Refrigeration* 63 (2016), pp. 184–198.
 - [32] Woohyun Kim and James E Braun. “Performance evaluation of a virtual refrigerant charge sensor”. In: *International Journal of Refrigeration* 36.3 (2013), pp. 1130–1141.
 - [33] Necati Kocyigit. “Fault and sensor error diagnostic strategies for a vapor compression refrigeration system by using fuzzy inference systems and artificial neural network”. In: *International Journal of Refrigeration* 50 (2015), pp. 69–79.
 - [34] T Lajos. “Áramlástan alapjai”. In: *Budapest: Műegyetemi kiadó* (2009).
 - [35] Béla Lantos. “Irányítási rendszerek elmélete és tervezése I”. In: *Egy változós szabályozások*. Akadémia kiadó, Budapest (2001).

- [36] Lars Finn Sloth Larsen. *Model based control of refrigeration systems*. Department of Control Engineering, Aalborg University, 2006.
- [37] Lars FS Larsen, Roozbeh Izadi-Zamanabadi, and Rafael Wisniewski. “Supermarket refrigeration system-benchmark for hybrid system control”. In: *2007 European Control Conference (ECC)*. IEEE. 2007, pp. 113–120.
- [38] Guannan Li et al. “An improved decision tree-based fault diagnosis method for practical variable refrigerant flow system using virtual sensor-based fault indicators”. In: *Applied Thermal Engineering* 129 (2018), pp. 1292–1303.
- [39] Guannan Li et al. “An improved fault detection method for incipient centrifugal chiller faults using the PCA-R-SVDD algorithm”. In: *Energy and Buildings* 116 (2016), pp. 104–113.
- [40] G Li et al. “Study on the support vector data description (SVDD)-based chiller sensor fault detection efficiencies”. In: *The 24th IIR International Congress of Refrigeration, Yokohama, Japan*. 2015.
- [41] Thomas M Ludden, Stuart L Beal, and Lewis B Sheiner. “Comparison of the Akaike Information Criterion, the Schwarz criterion and the F test as guides to model selection”. In: *Journal of pharmacokinetics and biopharmaceutics* 22.5 (1994), pp. 431–445.
- [42] Aersity M Lyapunov. “The General Problem of the Stability of Motion [in Russian]”. In: *Gostekhizdat, Moscow* (1950).
- [43] A Macmillan Neil and C Douglas Creelman. *Detection theory: a user's guide*. 2005.
- [44] Magdi S Mahmoud and Haris M Khalid. “Expectation maximization approach to data-based fault diagnostics”. In: *Information Sciences* 235 (2013), pp. 80–96.
- [45] Ashok Patel and Bart Kosko. “Optimal noise benefits in Neyman–Pearson and inequality-constrained statistical signal detection”. In: *IEEE Transactions on Signal Processing* 57.5 (2009), pp. 1655–1669.
- [46] Ralf LM Peeters, Bernard Hanzon, and Martine Olivi. “Canonical lossless state-space systems: Staircase forms and the Schur algorithm”. In: *Linear Algebra and its Applications* 425.2-3 (2007), pp. 404–433.
- [47] Linda R Petzold. *Description of DASSL: a differential/algebraic system solver*. Tech. rep. Sandia National Labs., Livermore, CA (USA), 1982.
- [48] Robert Gilmore Pontius and Benoit Parmentier. “Recommendations for using the relative operating characteristic (ROC)”. In: *Landscape ecology* 29.3 (2014), pp. 367–382.
- [49] Thomas Pursche et al. “Identification of Overtemperature Disturbances in Industrial Food Refrigeration Processes”. In: *2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*. IEEE. 2018, pp. 636–641.
- [50] Peng-Cheng Qi et al. “Experimental investigation of the optimal heat rejection pressure for a transcritical CO₂ heat pump water heater”. In: *Applied Thermal Engineering* 56.1-2 (2013), pp. 120–125.
- [51] Amro Quedan et al. “Low charge detection and diagnoses for non-critically charged refrigeration system”. In: *2017 IEEE International Conference on Electro Information Technology (EIT)*. IEEE. 2017, pp. 280–285.
- [52] David Rodriguez et al. “Parameter identification of a multi-stage, multi-load-demand experimental refrigeration plant”. In: *Control Engineering Practice* 60 (2017), pp. 133–147.
- [53] Ivan W Selesnick and C Sidney Burrus. “Generalized digital Butterworth filter design”. In: *IEEE Transactions on signal processing* 46.6 (1998), pp. 1688–1694.
- [54] Liang-Liang Shao, Zi-Yang Zhang, and Chun-Lu Zhang. “Constrained optimal high pressure equation of CO₂ transcritical cycle”. In: *Applied Thermal Engineering* 128 (2018), pp. 173–178.

-
- [55] Sigurd Skogestad and Ian Postlethwaite. *Multivariable feedback control: analysis and design*. Vol. 2. Wiley New York, 2007.
 - [56] Wanda Solano, Bikramjit Banerjee, and Landon Kraemer. “Particle Filtering for Model-Based Anomaly Detection in Sensor Networks”. In: (2012).
 - [57] Eduardo D Sontag. *Mathematical control theory: deterministic finite dimensional systems*. Vol. 6. Springer Science & Business Media, 2013.
 - [58] Shravan Srinivasan et al. “Bugs in the freezer: Detecting faults in supermarket refrigeration systems using energy signals”. In: *Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems*. ACM. 2015, pp. 101–110.
 - [59] Sylvain Verron, Teodor Tiplica, and Abdessamad Kobi. “Fault diagnosis of industrial systems by conditional Gaussian network including a distance rejection criterion”. In: *Engineering applications of artificial intelligence* 23.7 (2010), pp. 1229–1235.
 - [60] Etienne Videcoq, Daniel Petit, and André Piteau. “Experimental modelling and estimation of time varying thermal sources”. In: *International Journal of Thermal Sciences* 42.3 (2003), pp. 255–265.
 - [61] Zhanwei Wang et al. “Enhanced chiller fault detection using Bayesian network and principal component analysis”. In: *Applied Thermal Engineering* 141 (2018), pp. 898–905.
 - [62] Hong-Tzer Yang, Chao-Ming Huang, and Ching-Lien Huang. “Identification of ARMAX model for short term load forecasting: an evolutionary programming approach”. In: *Proceedings of Power Industry Computer Applications Conference*. IEEE. 1995, pp. 325–330.
 - [63] Zhenyu Yang et al. “Fault Detection and Isolation for a Supermarket Refrigeration System—Part One: Kalman-Filter-Based Methods”. In: *IFAC Proceedings Volumes* 44.1 (2011), pp. 13233–13238.
 - [64] Zhenyu Yang et al. “Fault Detection and Isolation for a Supermarket Refrigeration System—Part Two: Unknown-Input-Observer Method and Its Extension”. In: *IFAC Proceedings Volumes* 44.1 (2011), pp. 4238–4243.
 - [65] Kangkang Zhang et al. “Expectation–maximization approach to fault diagnosis with missing data”. In: *IEEE Transactions on Industrial Electronics* 62.2 (2014), pp. 1231–1240.

Appendices

A Conclusion and Perspectives of the Special Course

A.1 Conclusion

In this report, an entire, reliable model has been set up in Matlab for testing control methods for supermarket refrigeration systems. It provides better initial transients than an industrial alternative, and it models a larger scale of dynamics. However, the slowness of the simulations, the differences in time scales, and the higher potential to commit mistakes, when changing the code, gives a better trust to the TIL Suite library.

There were more achievements of the simulations. The different time scales in a thermodynamic system were simulated with a high fidelity, showing that although sound dynamics can be neglected, hydraulic resistance can actually speed up or slow down heat transfer. It turned out that modelling of the refrigerant loop cannot be done without controlled fans, since constant air temperature on the ambient air side lets the system unreasonably integrate energy (first, in the receiver). It also shows that the receiver is a good source of fault detection, being the buffer of the system.

In summary, the created model acts as a whole (it does not integrate energy, has reasonable time constants, and converges to the correct stationary values). Therefore model based state estimation or control design may be built upon it.

A.2 Perspectives

Modelling supermarket refrigeration systems is not done yet. There are missing popular elements, for example the heat recovery unit to heat up the supermarket, the air conditioning unit, a bypass valve of the gas cooler for cold days, in order to avoid the system cool too much, and the IT compressor, which saves on efficiency. Ejectors, as being the more efficient alternatives to high pressure valves, are also not considered, although they are not spread on the market yet.

The most important perspective is that different models are needed to answer different questions. This is why a good, general and robust library provides more use in the long term, than something specific and narrow. For example, an industrial problem of today, is to detect gas loop and recover from it efficiently. Gas loop is caused by an offset or hysteresis in the temperature measurement and results in circulating the gas in vain, while the evaporators are heating up. Finding a correct model for the diagnosis and the energy efficient way of avoiding (or handling) gas loop is an industrial motivation.

Since the slowness is partly caused by the high level usage of the CoolProp library, either low level usage, or parameter estimation may be considered, to achieve higher simulation speed. Furthermore, speed might be increased by implementing the implicit Matlab solver, although it is not as fast as the one provided by Dymola.

B Observability and controllability staircase forms

In this section, the controllability and observability staircase forms [46] are presented for a linearization point above the saturation curve.

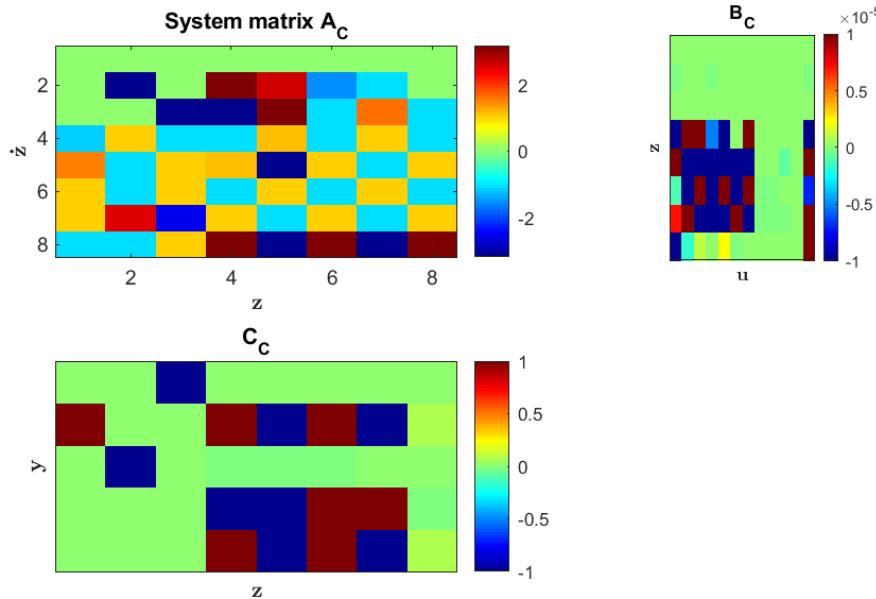


Figure B.1: Controllability staircase form

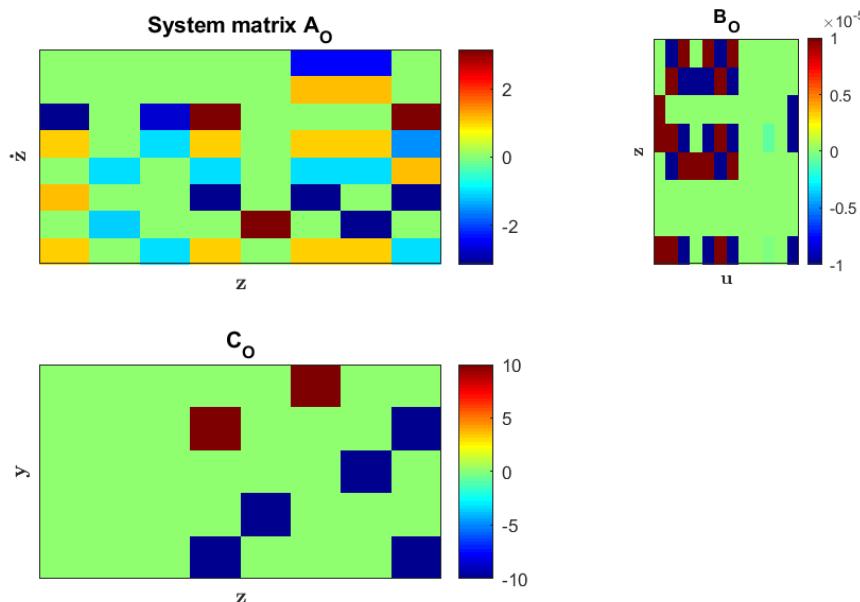


Figure B.2: Observability staircase form

C Simulink diagrams

In this section, the Simulink diagrams are presented, on which the simulations were carried out.

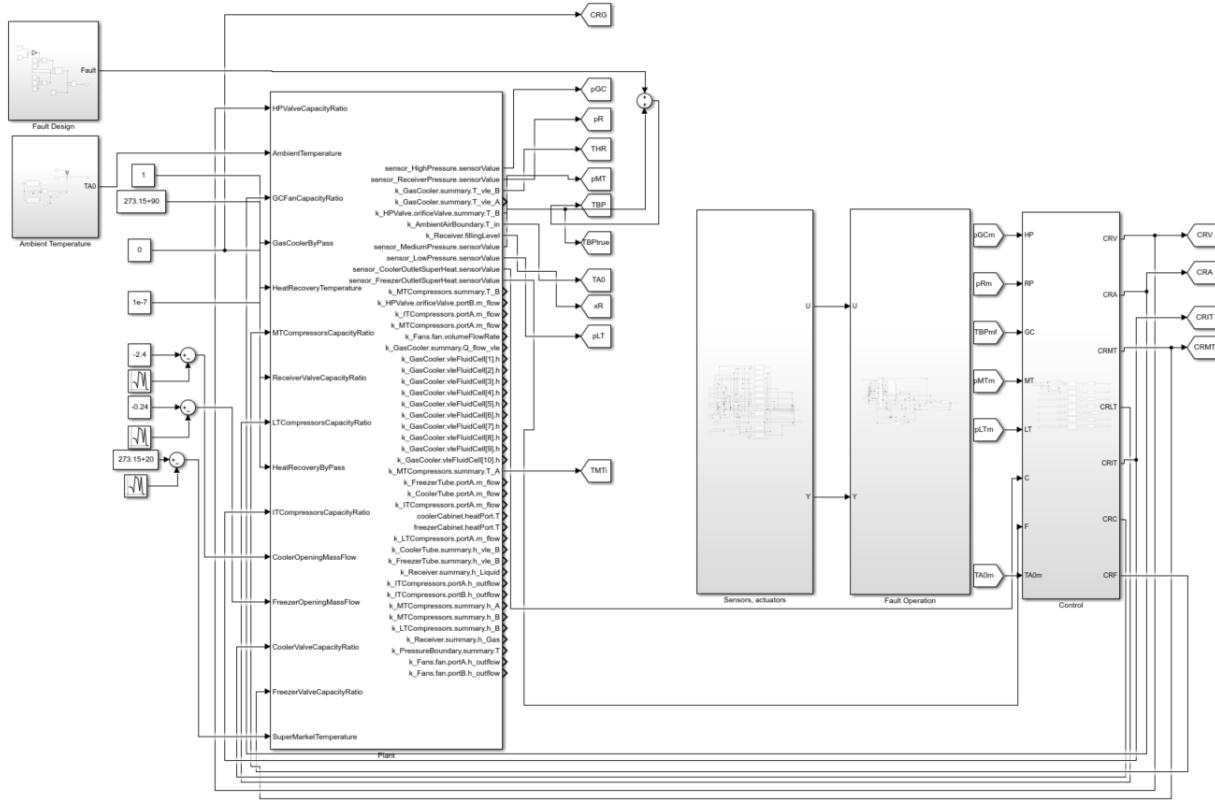


Figure C.3: Simulink diagram

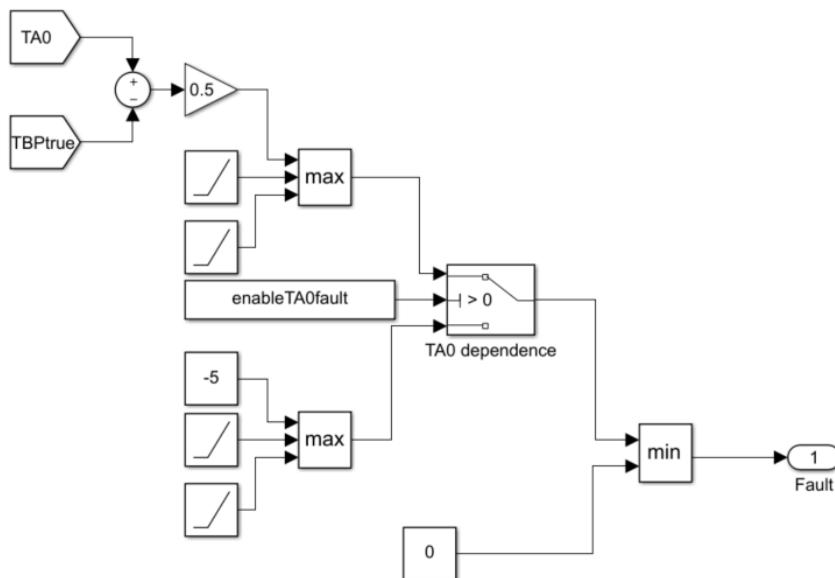


Figure C.4: Fault design

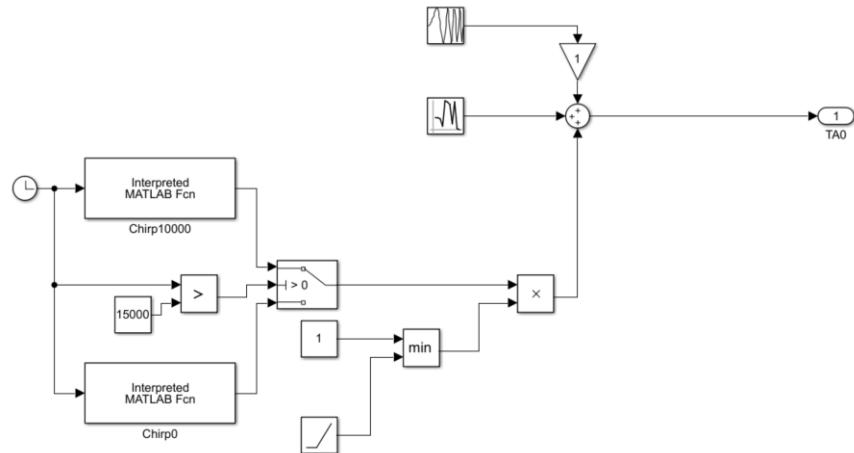


Figure C.5: Ambient temperature

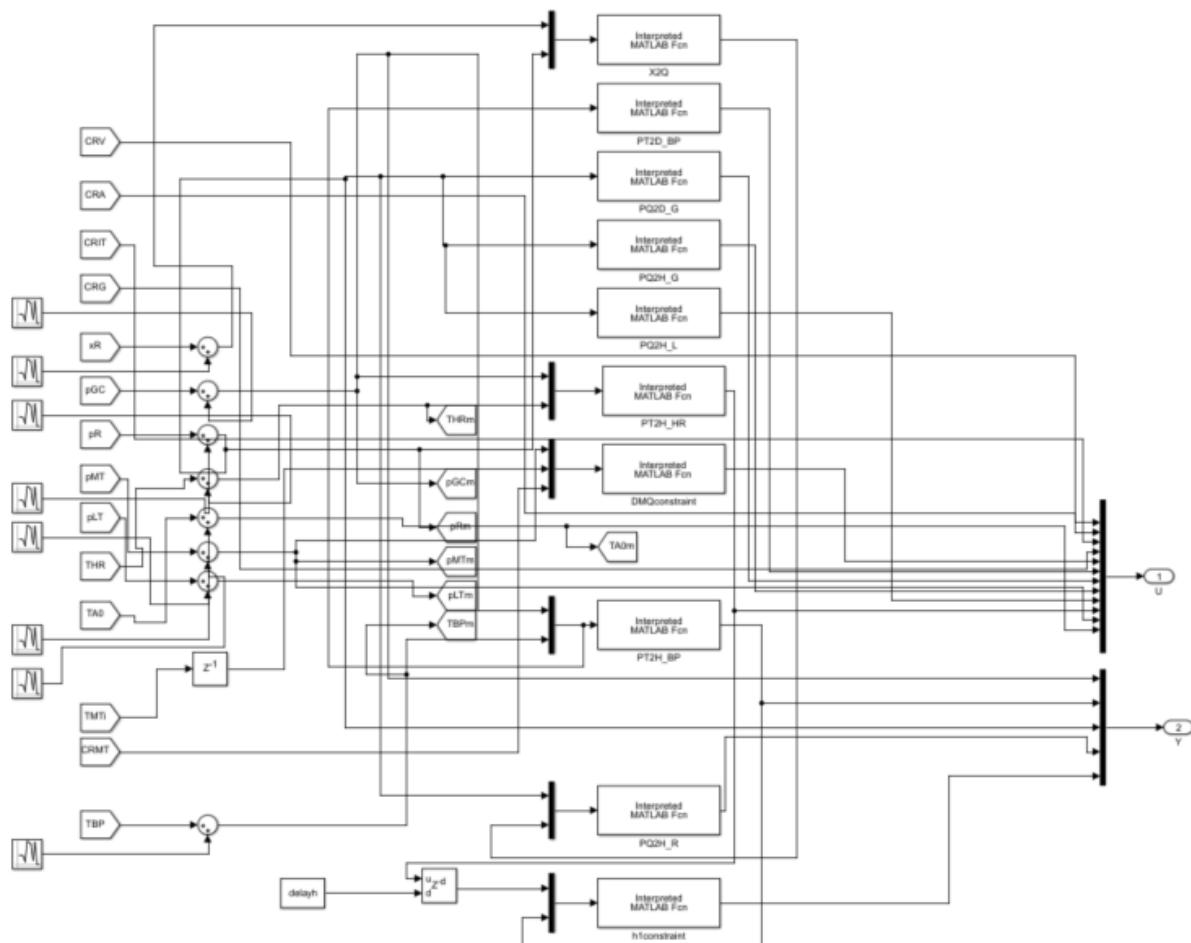


Figure C.6: Sensor emulator

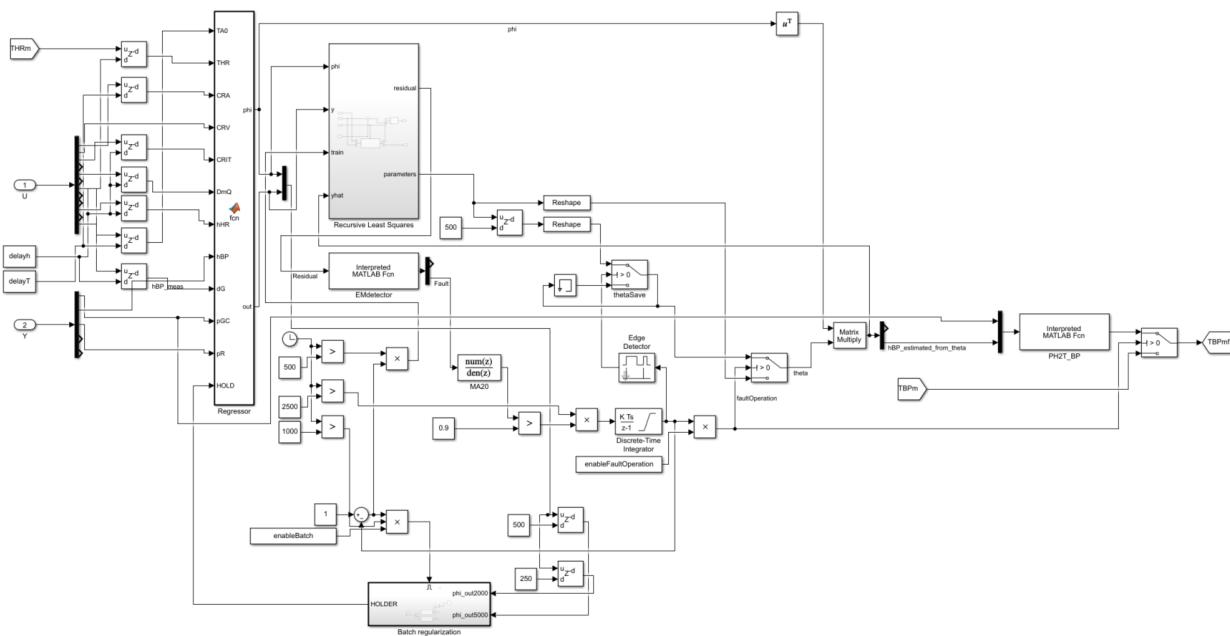


Figure C.7: Fault handling

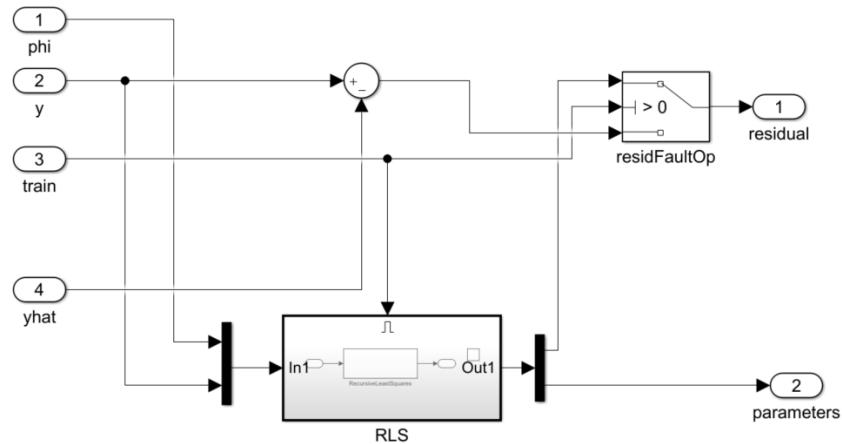


Figure C.8: Recursive linear least squares

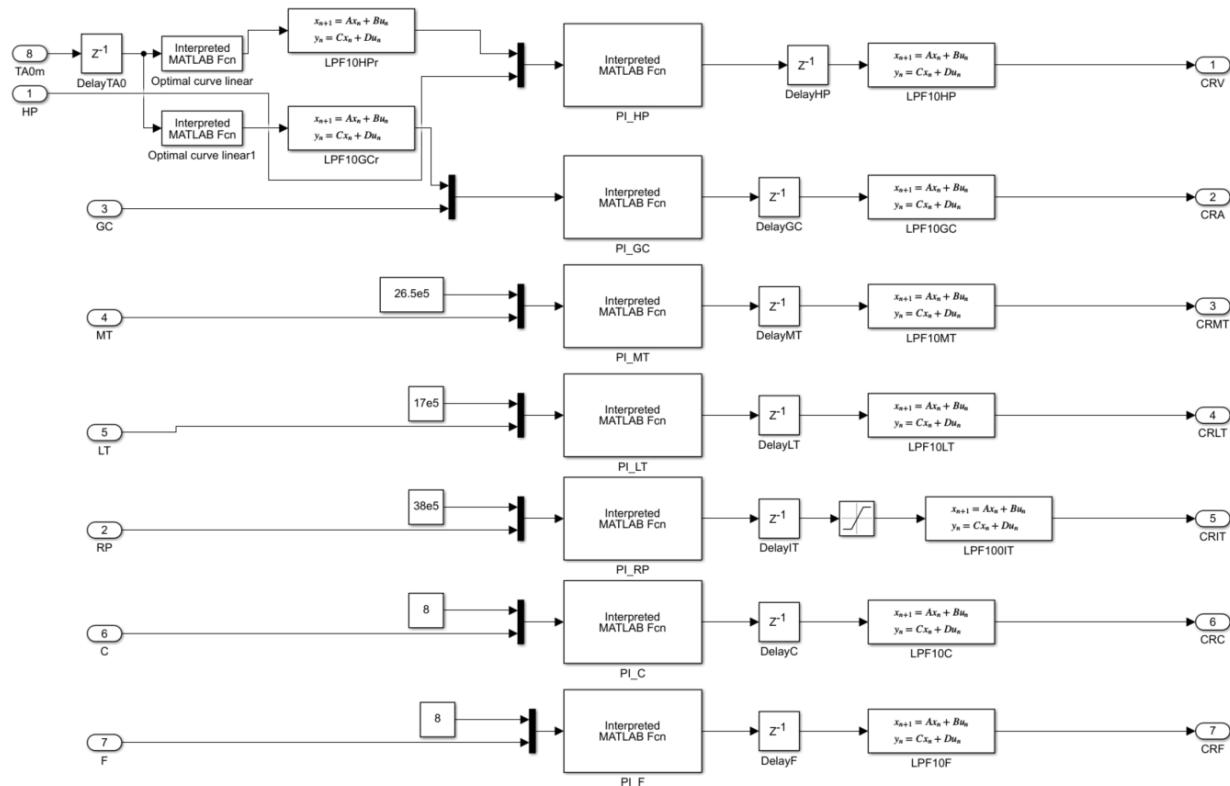


Figure C.9: Distributed controller

D Reference values and boundary conditions in the simulation

In this section, the reference values of the PI controllers, and the boundary conditions are presented.

Notation	Description	Value	Reference / Boundary condition
p_{GC}	High pressure	Linear fcn. of ambient temperature, 75...90 bar	ref. for HP valve (BP valve is closed)
T_{BP}	Gas cooler outlet temperature	Linear fcn. of ambient temperature, 30...33 °C	ref. for fans
p_{MT}	MT pressure	26.5 bar	ref. for MT comp.
p_{LT}	LT pressure	17 bar	ref. for LT comp.
p_R	Receiver pressure	38 bar	ref. for IT comp. (RP valve is closed)
sh_C	Cooler evaporator outlet superheat valve	8 K	ref. for cooler valve
sh_F	Freezer evaporator outlet superheat	8 K	ref. for freezer valve
$T_{A,0}$	Ambient temperature	Ranging between 10...30 °C	Boundary condition
$T_{SM,0}$	Supermarket temperature	20 °C	Boundary condition
$\dot{m}_{C,cab}$	Mass flow rate through cooler cabinet	2.4 kgs^{-1}	Boundary condition
$\dot{m}_{F,cab}$	Mass flow rate through freezer cabinet	0.24 kgs^{-1}	Boundary condition

Table D.1: Table of reference values and boundary conditions of the simulations

E Field experiment, back-end apparatus

In this section, a photo is depicted about the apparatus used for the field experiment.



Figure E.10: Photo of the field experiment back-end apparatus

F State estimation for the field experiment data

In this section, the results of the state estimation are presented, for the field data. They are presented here. Since the results could not be evaluated, like in the case of the simulation data.

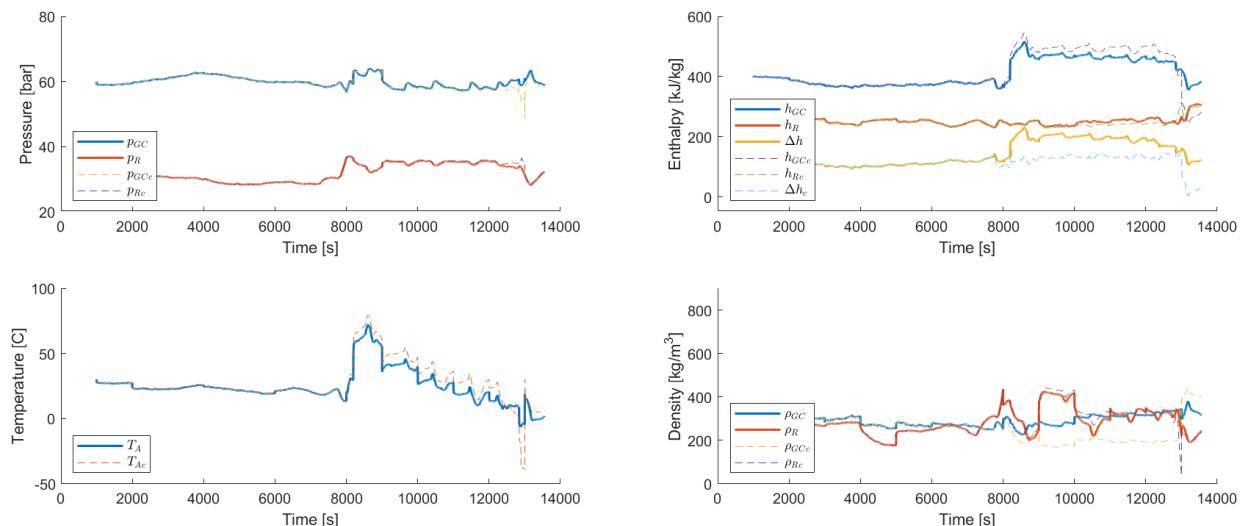


Figure F.11: State estimation for field data

DTU Electrical Engineering
Department of Electrical Engineering
Technical University of Denmark
Ørsteds Plads
Building 348
DK-2800 Kgs. Lyngby
Denmark
Tel: (+45) 45 25 38 00
www.elektro.dtu.dk