



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



Development of Machine Learning-based technologies for major appliances: soft sensing for drying technology applications

Ph.D. candidate
Giuliano Zambonin

Advisor
prof. Alessandro Beghi

Co-Advisor
prof. Gian Antonio Susto

Director & Coordinator
prof. Andrea Neviani

Ph.D. School in
Information Engineering

Department of
Information Engineering
University of Padova



**UNIVERSITÀ
DEGLI STUDI
DI PADOVA**

Head Office: Università degli Studi di Padova

Department

Department of Information Engineering

Ph.D. Course in: Information Engineering

Curriculum: Information and Communication technology

Series XXXII

Development of Machine Learning-based technologies for major appliances: soft sensing for drying technology applications

Thesis written with the financial contribution of Electrolux Italia S.p.A.

Coordinator: prof. Andrea Neviani

Supervisor: prof. Alessandro Beghi

Co-Advisor: prof. Gian Antonio Susto

Ph.D. student: Giuliano Zambonin

Summary

In this thesis, Machine Learning techniques for the improvements in performance of household major appliances are described. In particular, the focus is on drying technologies and domestic dryers are the machines of interest selected as case studies. Statistical models called Soft Sensors have been developed to provide estimates of quantities that are costly/time-consuming to measure in our applications using data which were available for other purposes.

This work deals with the following major applications areas:

1. Development of Soft Sensors for laundry moisture determination during drying cycles with the aim to improve the current automatic procedures in production.
2. Development of Soft Sensors for dry laundry weight determination during drying cycles with the aim to explore new procedures to improve the feature engineering phase involved in the workflow of the current algorithms.

The precise knowledge of the above-mentioned quantities may lead to more accurate drying, faster processed and lower energy consumption without increasing the production costs. Fabric care manufacturers are striving to make more energy-efficient and more user-friendly products. In interests of competitiveness, they are seeking to reduce both their electricity consumption and their drying time, possibly without increasing the hardware costs. The goal of this applied industrial research is to use innovative work methodologies such as the already mentioned statistical predictors to solve the above-mentioned problems.

Three major constraints characterize the development of Soft Sensors required in this thesis; they are summarized below and explained in section 1.5:

- i need for simple solutions for low-resourced hardware environment;
- ii limited availability of observations;
- iii need for interpretable models.

The following topics will be discussed here in details:

- Sparse regularization for variable selection: it was used for both the goals to provide simple and interpretable solutions using a limited amount of observations.
- Symbolic Regression: the evolutionary paradigm has been exploited to search the space of mathematical expressions to find typologies of models that best fit a given dataset avoiding the strong assumption of linearity. It was applied to the mentioned problem of laundry moisture estimation.
- A specific model has been chosen as a suitable structure to describe the laundry moisture evolution during drying cycles as a time series; parameters of such a model have been estimated in simulations at the beginning of the cycles to predict the laundry drying evolution. The resulting model is the one that has been proposed to improve the drying algorithms of machines as well as their calibration offline.
- Functional Regression techniques have been investigated in the problem of laundry weight estimation to exploit the total amount of the inherent information on available signals. This has allowed the possibility to study a benchmark for the performance and the development of a method for the determination of intervals of input signals that are most useful for the prediction goal improving the Soft Sensing feature engineering process.

The aforementioned works have been developed as industrially driven research activities during my industrial Ph.D. in collaborations with Electrolux Italia S.p.a. R&D department located in Porcia, Pordenone, Italy and Pontificia Universidade Católica do Paraná (PUC-PR), department of Production and Systems Engineering, Curitiba, Brazil. During the thesis, practical aspects of the implementation of the proposed approaches in a real industrial environment as well as topics related to collaborations between industry and academies will be specified.

Sommario

Nella tesi vengono descritti algoritmi di apprendimento automatico per migliorare le performance di macchine asciugabiancheria ad uso domestico. Sono stati sviluppate tipologie di modelli statistici denominati sensori virtuali allo scopo di fornire stime di quantità che non è possibile ottenere tramite misurazione diretta di sensori fisici o per le quali la determinazione tramite sensori fisici risulta onerosa; le suddette stime vengono ricavate sfruttando dati già disponibili per altre tipologie di progetto senza la necessità di effettuare ulteriori test di laboratorio *ad-hoc* per lo scopo oggetto di questa tesi.

Nello specifico, questo lavoro si articola nei seguenti temi:

1. Sviluppo di sensori virtuali per la stima del valore di umidità di biancheria durante i cicli di asciugatura allo scopo di migliorare l'attuale procedura automatica in produzione.
2. Sviluppo di sensori virtuali per la stima del peso di biancheria asciutta durante i cicli di asciugatura allo scopo di migliorare la fase di determinazione degli ingressi degli algoritmi attualmente utilizzati per questo obiettivo a partire dai segnali acquisiti.

La precisa determinazione delle quantità sopracitate può essere sfruttata per una più accurata gestione del ciclo di asciugatura e un ridotto consumo di energia senza incrementare i costi di produzione. I produttori di macchine asciugatrici e lavabiancheria ad uso domestico sono attualmente focalizzati nell'ottenere prodotti ad efficiente consumo energetico e di più semplice utilizzo. In un'ottica di competitività ci si concentra attualmente sulla riduzione dei tempi di asciugatura possibilmente senza incrementare i costi legati alla parte hardware delle macchine. Lo scopo di questo tipo di ricerca orientata ad applicazioni industriali è quello di usare metodologie di lavoro innovative come i già citati modelli di predizione statistica al fine di risolvere le problematiche sopracitate. Tre sono i principali vincoli che caratterizzano lo sviluppo dei sensori virtuali richiesti in questa tesi; essi vengono riassunti di seguito e spiegati nella sezione 1.5:

- i necessità di soluzioni semplici per un ambiente hardware ridotto;
- ii ridotta disponibilità di dati;
- iii necessità di modelli interpretabili.

I seguenti argomenti verranno discussi qui in dettaglio:

- Regolarizzazione sparsa per la selezione dei predittori: si è fatto uso della regolarizzazione per sviluppare soluzioni per entrambi i problemi citati in precedenza al fine di ottenere modelli di semplice implementazione e interpretabilità utilizzando un insieme di dati a numerosità ridotta.
- Regressione Simbolica: il paradigma di ottimizzazione ispirata ad algoritmi di tipo evuzionistico è stato sfruttato per ricercare tipologie di modelli in grado di effettuare un buon adattamento all'uscita disponibile senza la forte ipotesi iniziale di utilizzare strutture di tipo lineare. Si è fatto utilizzo di questo tipo di metodologie per il problema di stima dell'umidità di biancheria durante i cicli di asciugatura.
- E' stato scelto un modello specifico per la stima di umidità di biancheria descritta come serie temporale e i parametri di questo modello sono stati stimati in simulazione sfruttando le porzioni di segnali all'inizio dell' asciugatura al fine di ottenere una predizione del ciclo durante la sua prima fase. Tale modello è stato selezionato come alternativa all'attuale algoritmo di gestione del ciclo per il miglioramento delle performance e procedura di calibrazione offline.
- Sono state investigate tecniche di regressione funzionale per il problema di stima del peso asciutto di biancheria al fine di sfruttare l'intera informazione insita nei segnali disponibili. Tale studio ha reso possibile la determinazione delle finestre temporali dei segnali di ingresso che risultano più significative per le stime desiderate rendendo più agevole la fase di feature engineering fondamentale per lo sviluppo degli attuali algoritmi presenti in macchina per risolvere questo tipo di problema.

Le tematiche affrontate in questa tesi sono state sviluppate in accordo con un programma di ricerca industriale in collaborazione con il dipartimento di Ricerca & Sviluppo di Electrolux Italia S.p.a. di Porcia, Pordenone e in collaborazione con il gruppo di Ingegneria di Produzione e Sistemi dell'Università PUC-PR presso Curitiba, Paraná, Brasile. Verranno specificati nel proseguo della tesi gli aspetti pratici legati all'implementazione degli approcci proposti per la valutazione nello specifico ambito industriale assieme ai risultati derivanti dalla collaborazione tra realtà industriale e accademica, parte caratterizzante di questa attività di ricerca.

Preface

The research activities described in this thesis have been funded by Electrolux Italia S.p.a. and experimental activities were carried out at the Electrolux laboratories located in Porcia, Pordenone, Italy.

Contents

I	Modeling for fabric care household appliances	7
1	Introduction and Challenges	9
1.1	The importance of simulation modeling in industry: household appliances as a case study	9
1.2	Motivations of the thesis	12
1.3	Household clothes dryers: heat pump technology	13
1.4	Drying process: laundry moisture evolution	16
1.5	Fabric Care modeling guidelines (engineering constraints)	19
1.6	Objectives of the project	23
1.7	Contributions of this dissertation	24
1.8	Organization of dissertation	26
II	Methods and tools	29
2	Soft Sensing	31
2.1	Soft Sensors in the process industry	31
2.2	Model-driven and Data-driven Soft Sensors	38
2.3	Soft Sensors for fabric care applications	40
3	Algorithms for drying technology applications	43
3.1	Automatic End-of-Cycle determination	43
3.2	Time-to-End algorithm	51
3.3	Laundry weight determination	52
4	Methodologies	57
4.1	Multiple Linear Regression (MLR)	57

4.2	The Symbolic Regression method (SR)	67
4.3	Functional Data Analysis (FDA) overview	71
4.4	A priori hypotheses on research results	78
III	Soft Sensors for laundry moisture estimation	81
5	Dataset description	83
5.1	Heat-Pump Washer-Dryer dataset	83
5.2	Moisture reference determination	87
6	Physical models for moisture transfer in porous media	95
6.1	Drying phenomena	95
6.2	Moisture transfer by diffusion	96
7	Regularized multiple linear models	101
7.1	Description of the procedure	101
7.2	Experimental Setting	104
8	Symbolic models using Genetic Programming	107
8.1	Description of the procedure	107
8.2	Experimental setting	109
9	Polynomial model: laundry moisture prediction as a time series	113
9.1	Description of the procedure	113
9.2	Feature engineering	118
9.3	Experimental setting	120
10	Experimental results	123
10.1	Regularized linear models	123
10.2	Symbolic Regression	126
10.3	Polynomial model	128
10.4	Comparison between methods	130
IV	Soft Sensors for laundry weight estimation	137
11	Dataset description	139
11.1	Heat Pump Tumble Dryer data	139

12 Scalar-on-Function Regression models	143
12.1 Description of the procedure: Functional Linear Models (FLM)	143
13 Experimental results	149
13.1 Experimental settings	149
V Conclusions	157
14 Limitations and future works	159
14.1 Limitations of the research	159
14.2 Comparison with a priori hypotheses	161
14.3 Future works	163
15 Conclusions	165
References	169

List of Figures

1.1	Examples of internal tools used for data analysis, simulations and debug.	11
1.2	Washer-Dryer Heat Pump machine.	13
1.3	Representations of the main modules of a WD-HP machine.	15
1.4	Laundry moisture evolution during drying cycle.	17
1.5	Drying rate evolution during drying cycle.	18
1.6	The warehouse placed in Electrolux Porcia plant (PN), Italy.	19
1.7	Some Electrolux laboratories, R&D Fabric care department, Porcia, Por- denone, Italy.	22
1.8	Representation of the Soft Sensors development treated in this thesis explained in a nutshell.	25
1.9	Scheme of the dissertation by topics.	27
2.1	Soft Sensor in a control scheme.	32
2.2	A general methodology for Soft Sensor development.	34
2.3	Soft Sensor as a black-box model.	39
3.1	The Energy Label. Image taken from EL (2012).	44
3.2	The EoC signal components.	48
3.3	EoC signal with threshold for the cycle end detection.	49
3.4	EoC signal vs laundry moisture (Humidity) obtained from elaborations of the weight signal available from laboratory tests.	49
3.5	Inertia levels for different drying cycles. from left to right: small(A), medium(B,C) and large(D); nominal weights omitted.	50
3.6	User Interface, Electrolux Heat Pump Washer-Dryer (WD-HP).	51
3.7	Time-to-End (TTE) representation for drying tests with different laundry load size.	52
3.8	Laundry load estimation performed by the current algorithm on dryer machines.	54

3.9	Drying cycle handling, WD-HP case study.	55
4.1	Multiple Linear Regression (Soft Sensor).	59
4.2	Main situations in model fitting; from left to right: underfitting, appropriate, overfitting.	64
4.3	Sparse modeling visualization for Multiple Linear Regression problems. . .	66
4.4	Representation of GP-based Symbolic Regression.	67
4.5	The knowledge workflow in Symbolic Regression procedure.	68
4.6	Genetic Programming workflow.	68
4.7	Example of a tree to represent a symbolic expression.	71
4.8	An example of the estimated surface computed using a functional regression framework.	73
4.9	The trade-off between accuracy and interpretability of the required Soft Sensors.	77
5.1	Data visualization by (dry) laundry load weight: (a) Laundry load weight distribution in the data, 4 classes in [kg]; (b) laundry moisture reference.	86
5.2	Visualization of the correlation matrix (absolute values of correlation are reported): each row and column stands for an available signal.	87
5.3	The Experimental apparatus. WD machine on balance.	88
5.4	An example of weight acquisition from balance and relative offline elaboration to obtain the moisture reference.	91
5.5	Examples of visualization of the EoC signal.	92
5.6	The old and new reference. Above: laundry moisture during drying; below: the associated drying rate.	93
6.1	Particles randomly move through the porous medium according to a diffusion process.	97
7.1	Visual explanation of the algorithm. For each time sample an estimation value is available (e.g. dot depicted in each plot).	102
7.2	Multiple Linear Regression algorithm flow-chart.	103
7.3	Cost function used for the error evaluation in LASSO regularization. . . .	105
7.4	Visual Representation of the Multiple Linear Regression used for the laundry moisture estimation problem.	106
8.1	A multigene regression model population.	110
9.1	Visual explanation of the polynomial SS.	115

9.2	Algorithm flow-chart for the automatic EoC using polynomial model in time.	121
10.1	Examples of results using MLR.	125
10.2	Examples of results using Symbolic Regression.	127
10.3	Examples of results using the polynomial SS (poly3).	129
10.4	Laundry moisture SSs: comparison between methods in terms of estimation error distribution.	131
10.5	Q-Q plots.	135
11.1	Tests in the dataset divided by nominal weight.	140
11.2	The indication of drying remaining time (TTE) on a dryer machine. . . .	140
12.1	Interpretation of β -functions in FLM. Taken from Ogden and Goldsmith (2014).	145
13.1	Laundry weight SSs: comparison between methods in terms of the error distribution.	152
13.2	β functions obtained using FLM approach. Values on y-axis scaled in mean and standard deviation.	153
13.3	Q-Q plots associated to estimations made using each signal as unique predictor.	154
13.4	signal 2.	156

List of Tables

2.1	Problems with hardware sensors.	37
2.2	Purposes in using Soft Sensors.	37
2.3	Fabric care R&D projects based on SSs development.	40
4.1	Summary of the proposed methodologies for the design of the required Soft Sensor.	76
5.1	Dataset available for laundry moisture Soft Sensor. FMC stands for Final Moisture Target; it changes according to the dry program type.	85
9.1	Drying models fitted to experimental data; t =time, y =moisture.	117
9.2	Features used to estimate the parameters of the polynomial model.	119
10.1	Comparison between the proposed methods.	130

Acronyms

CV: Cross-Validation

DOE: Design of Experiments

EoC: End-of-the-Cycle

FDA: Functional Data Analysis

LASSO: Least Absolute Shrinkage and Selection Operator

ML: Machine Learning

MLR: Multiple Linear Regression

PFR: Penalized Functional Regression

RMSE: Root Mean Square Error

RR: Ridge Regression

SR: Symbolic Regression

SS: Soft Sensor

SSs: Soft Sensors

TD: Tumble-Dryer

TTE: Time-to-End

WD: Washer-Dryer

WD-HP: Washer-Dryer Heat Pump

Part I

Modeling for fabric care
household appliances

1

Introduction and Challenges

1.1 The importance of simulation modeling in industry: household appliances as a case study

Simulation development plays a major role in the industrial engineering field nowadays. The importance of simulation in industrial engineering was highlighted by the first pioneers in simulation research studies, since the 80s and it was considered among the first top three methodologies used by industrial engineers and operations researchers ([Banks \(1998\)](#)). Simulation studies have a central role in business and research and it is demonstrated by many research papers and empirical studies, whose number has been continuously increasing since the first conceptualization in the mid 80s. These results are also confirmed by recent surveys such as ([Negahban and Smith \(2014\)](#); [Mourtzis, Doukas, and Bernidaki \(2014\)](#)). Maybe the most notable and powerful definition of simulation modeling comes from ([Banks \(1998\)](#)) who state that simulation is the imitation of the operation of the real-world process or system over time. The importance of this statement is confirmed by their recently published papers (for example [Funke and Jonsson \(2019\)](#)), and so it is taken as a reference in the rest of the thesis.

In general, the main benefits of Simulation-Driven Design (SDD) reported in the literature (e.g. [Shephard, Beall, O'bara, and Webster \(2004\)](#); [Thomke \(1998\)](#)) and shared in this

dissertation could be summarized in:

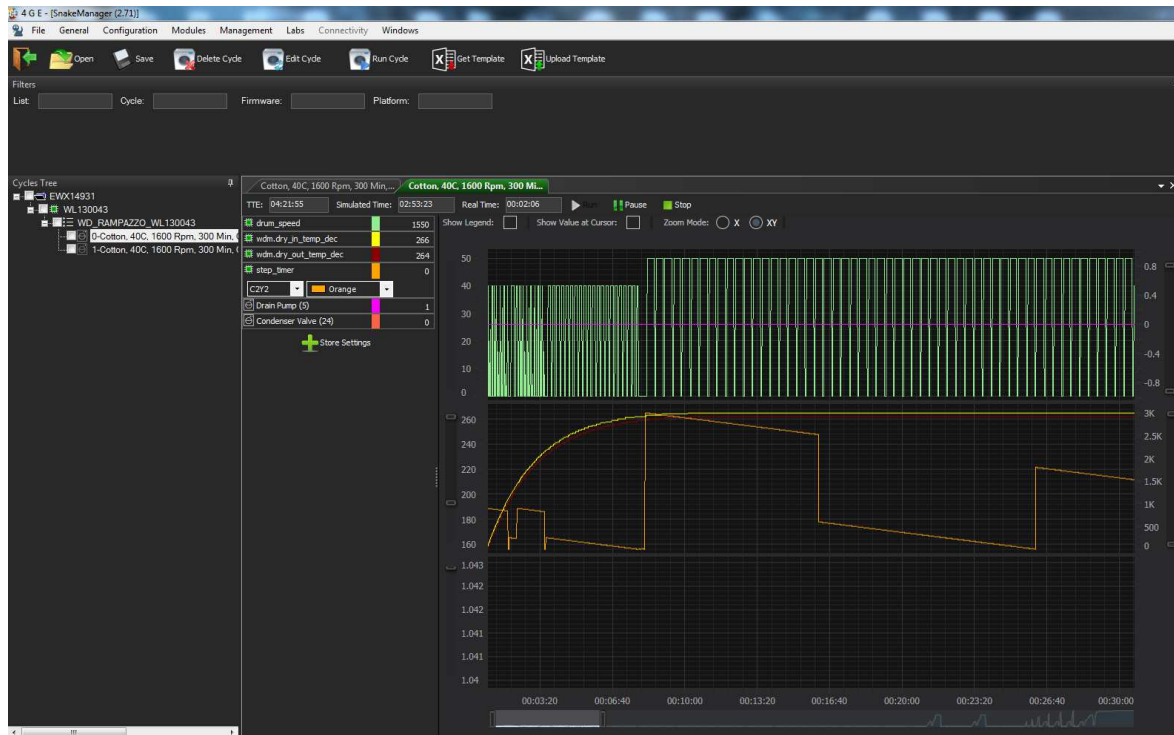
- decreased lead time and time-to-market;
- lower design costs;
- increased quality of products.

Furthermore, as stated in [Becker, Salvatore, and Zirpoli \(2005\)](#), virtual simulation tools do not only decrease development time and cut costs, but also alter the way engineers solve problems. It is argued that when using physical testing, mostly conservative innovation is achieved because previous generations of designs more often will act as a basis for the new. By using virtual simulation tools the engineers can observe phenomena in a much more controlled environment, and therefore develop new problem-solving abilities that enables more radical innovation. Dealing with challenges to Simulation-Driven devices, the main problems were identified in a study conducted by [Abe \(2006\)](#) when companies are understaffed, overburdening the design engineers with work. Indeed, this creates further issues as they do not feel as they have the time to learn how to simulate, which in turn would reduce the lead time and their workload.

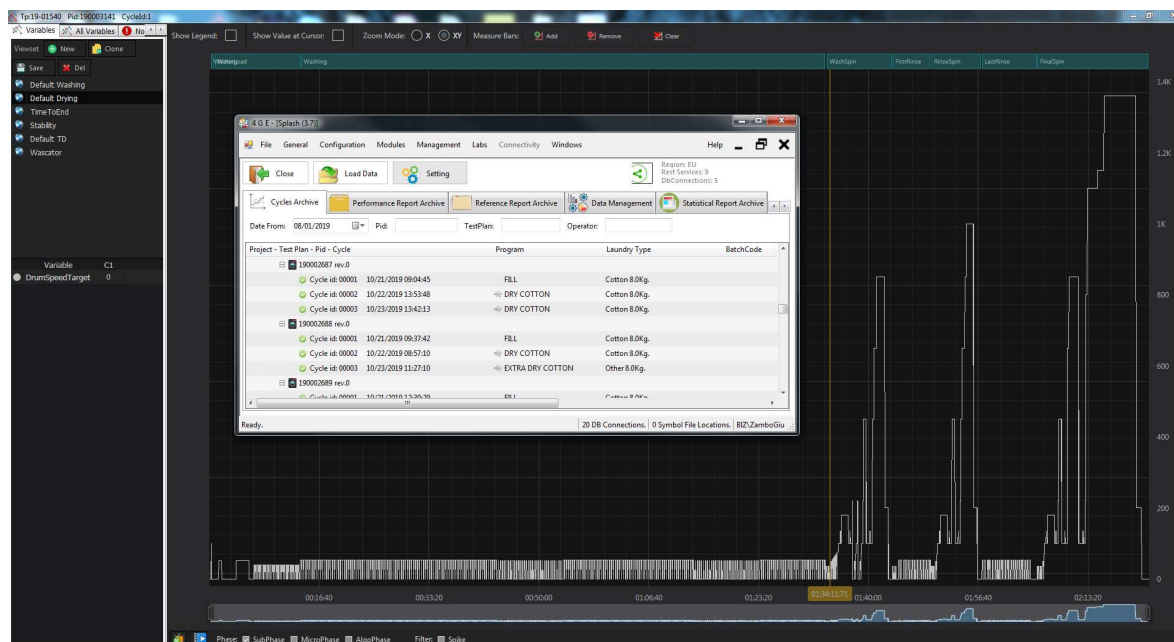
In the case study of this dissertation, the entire work aims to improve the current frameworks exploited for SDD to obtain models able to predict the evolution of the process of interest (in this dissertation drying cycles in household appliances will be treated as the case study) making possible the evaluation of the best controls for the cycles on machines. Simulation-Driven devices represent crucial instruments for fabric care projects treated in this thesis and for several reasons; in fact, they are constantly and intensively used by technicians to carry out the debug of the software designed for specific washing/drying cycles and to develop new strategies for the process handling. In [Figure 1.1](#) a representation of the main simulation tools currently used in Electrolux fabric care R&D is depicted; the improvement of the features of such a tool is ongoing with the aim to provide increasingly reliable instruments to simulate real processes using software and avoiding the execution of time-consuming physical washing/drying cycles. The research activity carried out in last years and discussed in this dissertation, is focused on supplying new potential instruments and models to enrich the current set of features available for simulation devices and to provide new algorithms for the improvement of the online performance of cycle handling and their calibration phase.

1.1 The importance of simulation modeling in industry: household appliances as a case study

11



(a) The tool used for simulations of the cycles and debug.



(b) Another tool used for data analysis.

Figure 1.1: Examples of internal tools used for data analysis, simulations and debug.

1.2 Motivations of the thesis

In this thesis the main focus is on the proposal of soft sensors (or virtual sensors) for applications on fabric care household major appliances. Contributions of simulation modeling in this domain are reviewed and some original works have been produced.

The final goal of this dissertation is to prove, through applications, that the use of statistical modeling algorithms may constitute possible solutions even for industrial problems where lack of available data, significant restrictions on computation burden and difficulties on the physical description of the process are the main constraints.

In this introductory part of the thesis, the features of the research project are described in detail together with the fabric care modeling guidelines which are of primary importance for all activities in this industrial domain. Drying technology applications will be the objects of the entire dissertation even if some methods are flexible and adaptable for washing technologies and it will be expressly stated in such a case.

The two problems discussed are related to drying cycles in household dryers machines (Washer-Dryers and Tumble Dryers) and they deal with the estimation of two fundamentals quantities for the control of a drying cycle:

1. laundry moisture content;
2. laundry (dry) weight.

A detailed explanation of these two goals will be provided in the following together with details about their importance for the improvement of the current performance of the cycles implemented in the machines on the market.

This thesis discusses the main difficulties linked to the development of soft sensors for fabric care applications with the analysis of several methodologies for the estimation of the quantities of interest. Contributions of this type of research are therefore addressed to the following results:

- proposal of different flexible estimation techniques to achieve knowledge of the unknown quantities of interest;
- proposal of a new online algorithm for the automatic end-of-the-cycle. Possible benefits in terms of drying performance, calibration and cost reduction of such a new method will be discussed.

Another topic developed in this section is the description of the main modules of the machines (heat pump and condense technologies in particular) in order to convey the complexity of the physical process of drying fabrics with household appliances.

1.3 Household clothes dryers: heat pump technology

As already stated, drying cycles of Washer-Dryer machines are the focus for methodologies and algorithms developed in this dissertation; in particular, the heat pump technology has been taken into account because it is currently widespread in the European market. Heat pump dryers are closed-cycle devices that use a heat pump process to heat the air. This technology is characterized by low energy consumption and low temperature of the drying air, which makes the drying even more gentle for any type of fabrics and colors. This thesis, as already stated, is focused on this kind of washer-dryer and from now on we'll refer to this category only with the emphasis on heat pump modules of the machine which will be discussed below.

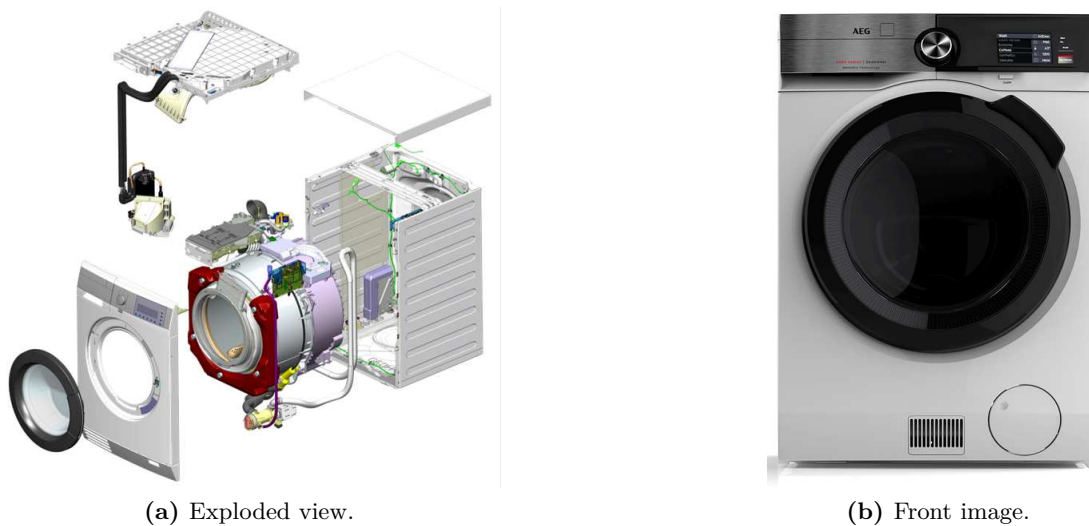


Figure 1.2: Washer-Dryer Heat Pump machine.

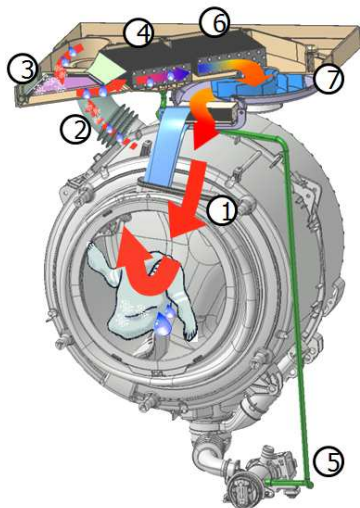
As anticipated, the Washer Dryer-Heat Pump (WD-HP) works in a closed system, resulting in a much more efficient process. In Figure 1.2 we have a representation of the machine of interest is depicted while in Figure 1.3 we can see details about the core part of the machine together with the heat pump modules. Describing in brief how this type of machine works: the fan extracts the moistened air from the tumble drum and forces it through the evaporator, removing the moisture; then the air stream is driven through the condenser, before re-entering the drum. Thus, the heat pump refrigerating capacity is used to dehumidify the moistened air, whereas the heating capacity is used for warming up the cold air. In particular: referring to Figure 1.3a:

1. the dry air enters the drum and absorbs moist from clothes;

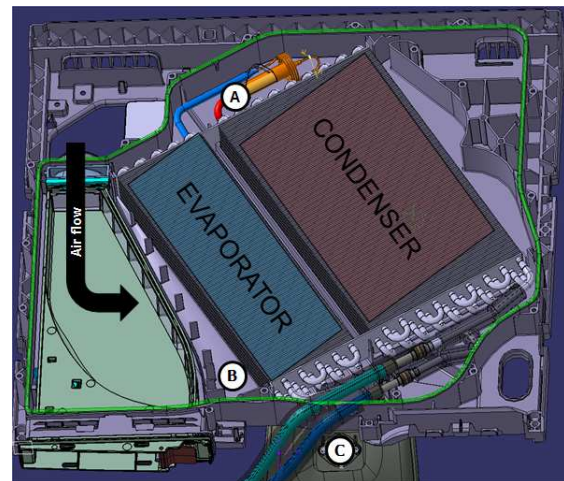
2. the moist air leaves the drum and enters the heat pump module (Figure 1.3b);
3. the filter retains most of the fluff released in air by clothes minimizing heat exchangers clogging;
4. the evaporator cools down air condensing moist;
5. the condensed moist is directed to the pump body and discharged;
6. the condenser heats dry air;
7. fan generates air flow circulating in the air duct (Figure 1.3c) and through the air filter (Figure 1.3b).

The process is allowed by the evaporating, compressing, condensing and expanding of a refrigerant: when the drying cycle begins, the gas in the circuit is conveyed by the compressor (Figure 1.3d) to the condenser, where at the outlet the circulation pipe known as capillary narrows. This capillary slows down the release of gas and ensures that the pressure inside the condenser increases bringing the gas to a liquid status; this change permits the proper functioning of the condenser and the evaporator in the heat pump module (Figure 1.3b). The heat pump refrigerating capacity is used to dehumidify the moistened air and the heating capacity is used to warm up the cold air. In this way part of the energy of the air that exits the drum (in form of heat and steam) is recovered and used to heat again the air for the following stream, causing a lower energy consumption compared to resistance air-heating based machines.

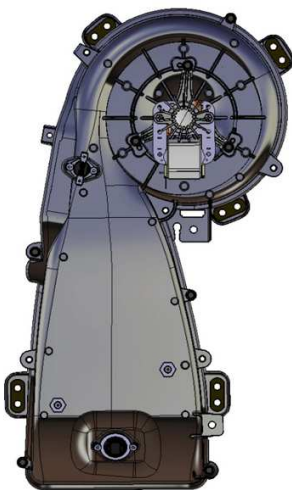
Figure 1.3b shows also the positions of the Negative Temperature Coefficient (NTC) sensors, respectively: **A** to measure the temperature of a capillary of Heat Pump module, **B** to measure the hot air temperature and **C** to measure the wet air temperature. In the next sections of the dissertation, temperature measures from these sensors will be processed and exploited as inputs for the design of Soft Sensors. As discussed in the following, these NTC together with the motor torque evolution represents the main source of information for the current automatic drying algorithm as well.



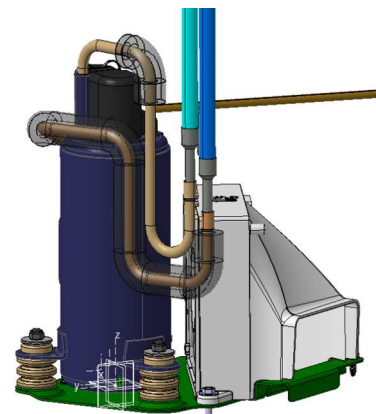
(a) The air cycle.



(b) Heat pump modules; A, B and C are the positions for NTC sensors.



(c) Air duct.



(d) Compressor module.

Figure 1.3: Representations of the main modules of a WD-HP machine.

1.4 Drying process: laundry moisture evolution

Basically speaking, drying can be described as a thermally driven process. Therefore, during drying, at least two main processes occur simultaneously: heat transfer from the surrounding environment toward the moist material to evaporate the moisture at material surface and transfer of internal moisture to the surface of the solid. Other processes may occur in some cases, namely, physical and chemical transformations which may induce changes in the structure of solids and fluids involved. As shown in [Dinger and Zamfirescu \(2016\)](#) there appear to be four major modes of transfer: (i) capillary flow of moisture in small interstices, (ii) moisture diffusion due to concentration gradients, (iii) vapor diffusion due to partial pressure gradients, and (iv) diffusion in liquid layers adsorbed at solid interfaces.

The mechanisms of diffusion will be treated in detail in the next sections as well as the identification of fundamental methods of analysis and prediction of the drying process of porous materials. Another important aspect is the mathematical modeling of the drying processes and equipment. Its purpose is to allow design engineers to choose the most suitable operating conditions and then size the drying equipment and drying chamber accordingly to meet desired operating conditions. The principle of modeling is based on formulating a set of mathematical equations for the process and on the assumption of boundary and initial conditions, which can adequately characterize the system. The solution of these equations must allow for the prediction of the process parameters as a function of time at any point in the drying equipment. The shrinkage, the porosity of porous materials, and the heat and moisture transfer analyses of drying in porous materials were also analyzed.

In a drying process, the moisture content is generally determined on a dry material basis. More exactly, when the weights of the moist and dry materials are compared, the mass of moisture is determined. The ratio between the mass of moisture and the mass of dry material represents the moisture content in percent on a dry basis; this is also denoted with the symbol y . The variation of dry basis moisture content is usually measured versus the drying time. Figure 1.4 shows the typical variation of moisture content during drying for a general case when a moist solid loses moisture. During the drying of a moist solid in heated air, the air supplies the necessary sensible and latent heat of evaporation to the moisture and also acts as a carrier gas for the removal of the water vapor formed from the vicinity of the evaporating surface.

In the diagram from Figure 1.4, the first part of the process, represented by curve A–B, occurs by mass transfer from the solid surface. This is a stage of warming up of the solid(s) during which the solid surface conditions come into equilibrium with the drying

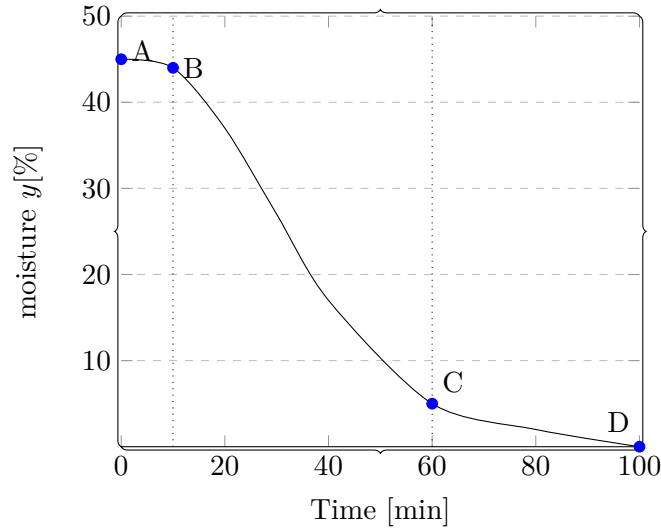


Figure 1.4: Laundry moisture evolution during drying cycle.

air. In terms of energy consumption, period A–B often represents a negligible proportion of the overall drying cycle requirements. During this period, the moist material can even take more humidity instead of being dried, depending on the actual conditions. Eventually, the surface reaches such conditions that the humidity diffuses out of the product.

The process B–C shows a reduction of the moisture content which is approximately linear in time. During this phase, the area of the saturated surface decreases gradually. This is the period of drying during which the rate of water removal per unit of drying surface is essentially constant. Point C, where the constant rate period ends, is known as the point of critical moisture content. During B–C period, the movement of moisture within the solid is rapid enough to keep a saturated condition at the surface, and the drying rate is controlled by the rate at which heat is transferred to the evaporating surface. The surface of the solid remains saturated with liquid water (or moisture) because movement of water within the solid to the surface takes place at a rate as great as the rate of evaporation from the surface. This stage is controlled by the heat and/or moisture transfer coefficients, the area exposed to the drying medium, and the difference in temperature and relative humidity between the drying air and the wet surface of the solid.

Once the saturated moisture at the surface is completely eliminated, the phase C–D follows during which there is no evaporation at the surface, but rather a diffusion process of the moisture within the solid followed by a convective mass transfer at the solid surface. From point C onward, the surface temperature begins to rise and continues to do so

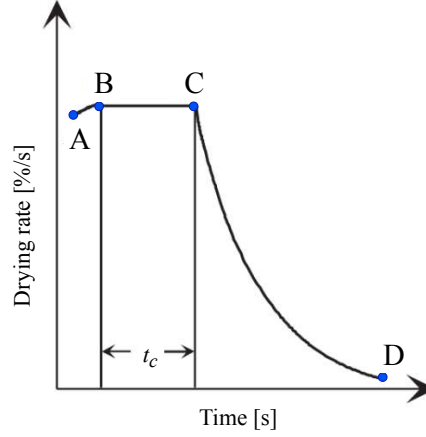


Figure 1.5: Drying rate evolution during drying cycle.

as drying proceeds, approaching the dry-bulb temperature of the air as the material approaches dryness. Therefore, when the initial moisture content is above the critical moisture content, the entire drying process occurs under the constant rate conditions. If it is below the critical moisture content, the entire drying process occurs in the falling rate period. A more convenient visualization of the drying phases is the drying rate which is the derivative of the moisture curve in time. An example of such visualization is represented in Figure 1.5 which represents the discussed main steps. t_c is the interval in which the rate of water removal per unit of drying surface is approximately constant. This type of visualization is appropriate to identify the drying cycle evolution and especially the last part of interest for the automatic detection of the end of the cycle that is one of the topics of this thesis. The drying rate curve visualization was used in this research for the reconstruction of the moisture reference as will be explained in detail in the following parts of the dissertation.

Drying kinetics refers to the changes of average material moisture content and average temperature with time, unlike drying dynamics which describes changes in the temperature and moisture profiles throughout the drying body. Drying kinetics enables calculating the amount of moisture evaporated, drying time, energy consumption, and so on. The changes in material moisture content and temperature are usually controlled by heat and moisture transfer between the body surface, the surroundings, and the internal structure of the drying material. The change in moisture content with time is influenced significantly by the parameters of the drying process, for example, temperature, humidity (pressure), the relative velocity of air, or total pressure.



Figure 1.6: The warehouse placed in Electrolux Porcia plant (PN), Italy.

1.5 Fabric Care modeling guidelines (engineering constraints)

In 1.1 we’ve already talked about the importance of simulation modeling in industrial contexts. Here the focus much more on data-driven modeling approaches for industrial applications and in particular for the products which are the case study of this thesis, i.e. household major appliances.

In Solomatine and Ostfeld (2008) data-driven modelling (DDM) can thus be considered as an approach to modelling that focuses on using the ML methods in building models that would complement or replace the “knowledge-driven” models describing physical behavior. DDM uses the methods developed in the fields mentioned above, and the role of a modeller is to tune them to a particular application area. *Modelling* in the name stresses the fact that this activity is close in its objectives to traditional approaches to modelling, and follows the traditionally accepted modelling steps, and that it does not comprise the analysis or mining of data only. The main part of data-driven modelling is, in fact, learning which incorporates the so-far unknown mappings (or dependencies) between a system’s inputs and its outputs from the available data (Mitchell (1997), Figure 2.1). Dealing with the specific industrial applications of interest: fabric care major appliances research and development; some properties of the projects related to the development of new models for the software modules of the machines, could be recognized. Projects of this type are sometimes aimed at the development of soft sensors (Data-driven models) for the determination of an unknown quantity useful for the cycle handling exploiting available information on machines. Considering these goals and requirements I’ve seen in

these years, I summarize and discuss them starting with the list below:

1. need for simple solutions for low-resourced hardware environment;
2. limited availability of observations (cycle acquisitions);
3. need for interpretable models (for debug and calibration phases).

In particular, point 1 expresses a strong requirement for the design of the above-mentioned soft sensors and it is very linked to the physical restrictions about the computational effort that characterizes the machines on the market. Whenever a specific algorithm is designed to fulfill a task on machine during cycles it is always necessary to evaluate the firmware solution for such a task if the studied solution requires too much memory occupancy or accuracy demand. For this reason, even if a procedure appears to be suitable for the problem at hand, its complexity is always an essential discriminator between other different possibilities, if available.

Point 2 is the one that most characterizes the projects discussed in this dissertation and represents the main constraint of the entire modeling design procedure for the problems at hand. Practically speaking, in the case of drying technology applications, collecting a number of complete drying cycles from laboratory is always time-consuming because a single test takes hours to get finished and the complexity of on-going projects together with their priority for the production makes the gathering of tests in exactly the same conditions for the same goal very difficult. That's the reason why already available acquisitions were exploited here for the development of soft sensors in this thesis, in order to avoid an excessive outflow of time and cost resources. So, it's always hard to see hundreds of tests available for data analysis and design of black-box models. Figure 1.7 shows images of Electrolux laboratories where data used and discussed in this thesis were collected.

The data scarcity problem discussed in this dissertation is therefore at the same time a limitation and a requirement for soft sensor design here, because whenever a new procedure is ready for implementation, it's always necessary to take into account the calibration phase which involves the adaptation of such a procedure to different cases (laundry typologies) and goals (e.g. energy saving).

Point 3 is naturally linked to the previous ones for what said and it's emphasized here to clarify the typologies of models (soft sensors) that are targets for this thesis whenever a new algorithm is requested for the online implementation on machines. So, an essential

requirement deals with the use of interpretable models, i.e. by simplifying, models for which the processing of inputs to obtain outputs is always clear and fixed. This point will be discussed in details in the course of the thesis when the use of several methodologies will be justified (Parts [III](#), [IV](#)).

As regards the definition of the concept of interpretability, as stated in [Molnar \(2019\)](#), there is actually no mathematical definition of interpretability. A (non-mathematical) definition from [Miller \(2018\)](#) is: interpretability is the degree to which a human can understand the cause of a decision. Practically speaking, the higher the interpretability of a machine learning model, the easier it is for someone to comprehend why certain decisions or predictions have been made; a model is better interpretable than another model if its decisions are easier for a human to comprehend than decisions from the other model. An interpretation of an erroneous prediction helps to understand the cause of the error. It delivers a direction for how to fix the system.



(a) Drying technology laboratory.



(b) Washing technology laboratory.

Figure 1.7: Some Electrolux laboratories, R&D Fabric care department, Porcia, Pordenone, Italy.

1.6 Objectives of the project

In the previous section, general guidelines for the design of data-driven models for applications in fabric care machines were provided; here the objective is to make clear the challenges of the particular problems discussed in this thesis concerning models for drying technology mentioned above and associated to the previous points.

As stated before, the two estimation problems discussed in this dissertation are related to the following two variables: I) laundry moisture content and II) laundry weight, both of them inferred during drying cycles using already available signals. The challenges inside the design of models for the determination of these quantities can be summarized taking the 3 constraints explained in the previous paragraph and considering two additional ones:

I the lack of humidity sensor on WD-HP machines;

II the time needed to make estimates available online on machines: estimates must be made available as soon as possible during drying cycles.

Point I regards the use of a specific machine as the case study for this thesis: indeed, data used for the development of soft sensors for moisture come from a heat pump washer dryer which does not have a physical dedicated sensor for the determination of moisture evolution during drying like tumble dryers have. The absence of such a sensor on wd-hp is due to limitations related to the technology and costs: the machine is a washer-dryer so it can be used to wash the laundry and/or dry it later and the presence of moving water inside the drum during washing cycles makes the use of cheap laundry moisture sensors problematic. This is actually the main reason to justify the choice of this case study for the research: try to use already available signals on wd-hp to infer through virtual sensors the quantity that is determined by physical dedicated sensors in domestic tumble dryers.

Point II is shared by the two estimation problems and is related to the request to be precise in the determination of the unknown quantities at the beginning of the cycle (first minutes) where it is known that the precision goal is hard taking into account the variability of the laundry composition and operating conditions.

All these challenges make the reader conscious about difficulties implied in the process of design of data-driven models for fabric care household major appliances and they will be recalled often during the dissertation in order to justify several methodological choices. The research activity described in this thesis is based and focused on industrial requirements and, in particular, it was specifically promoted and adapted to the needs of the industrial partner engaged in the improvement of the performance of products together

with the energy savings and customer satisfaction. For these purposes, my research activities held in R&D department during these years have been concentrated on two main goals (or specific objectives):

1. the development of new models (SSs) for the improvement of drying cycle handling and performances reducing the calibration effort which is the main current drawback of the procedure already available on machines;
2. provide new tools, methodologies and reports to the laboratories involved in the daily work with machines and dedicated on one side to the improvement of drying performance (time and energy savings, care of the product), on the other side strongly focused on cost reduction of the components to maintain the competitiveness on the market.

1.7 Contributions of this dissertation

Taking into account the aforementioned specific objectives, the contribution of the entire work discussed in this dissertation is threefold:

- (i) it provides an example of applied research in the design of Soft Sensors oriented to the daily consumer goods such as major appliances; few works are available in the literature in this peculiar field;
- (ii) one of the main results is the comprehension of what kind of model is more suitable to describe the laundry drying assuming only the use of data-driven approaches;
- (iii) it promotes the use of a functional regression framework to address the laundry weight estimation problem which has been treated in the literature using dedicated physical sensors, or multiple linear regression with hand-craft feature extraction. This thesis encourages the use of functional methods to improve both the performance and the interpretability of the provided data-driven models.

As regards the first point, two types of Soft Sensors are promoted in this work, one for each problem of interest: laundry moisture prediction and laundry nominal weight estimation. The common framework in which these SSs have been constructed is summarized in Figure 1.8; the two problems are different although associated with the same project phases. The design of such SSs provide significant information to set controls for laundry drying process which is a process characterized by extremely variable agents such as the composition of laundry, the initial moisture amount, the laundry movement inside the



Figure 1.8: Representation of the Soft Sensors development treated in this thesis explained in a nutshell.

drum.

This research on one side has the purpose of finding innovative solutions referring to the current set of implemented algorithms and for this reason, it considers the state of the art functional methodologies in the literature to perform the feature engineering process. On the other side, it gives also a priority to the needs of laboratory experts and engineers focused on interpretable solutions characterized by minimal efforts in the calibration process. Shortcomings will be claimed in the core part of the dissertation and with the aim of being as clear as possible they are summarized here for the reader:

- the results presented, i.e., the Soft Sensors (SSs) for the two main problems of interests, come from offline simulations working with typical drying process conditions of variability. Online test with a preliminary firmware version of the proposed solutions are planned but not available here for discussion;
- the use of different approaches for the design of the above-mentioned SSs, typically methodologies which exploit huge amounts of data (for example deep learning), are not treated in this thesis. Even though for the laundry weight estimation task more

tests could be available for offline analysis, these models are only mentioned here considering that their use could be useful to develop suitable procedure available for laboratories for cycle debugging and to simulate real cycle integrating existing tools (such as the one represented in Figure 1.1a).

The first part of the dissertation is dedicated to the explanation of the research question and requirements, presenting also the initial expectations for results based on intuition and the limited physical knowledge of the process. The other parts will be held to present and comment on the obtained results discussing if such results match the initial expectations or not raising hypotheses on reasons.

1.8 Organization of dissertation

This dissertation is organized as follows. In Part II characteristics of Soft Sensors for industrial applications are analyzed in detail as well as the main methodologies related to Regression problems pointed out in the course of the thesis. In this part the reader can also find the formalization of the goals considering their differences with the current solutions for the automatic handling of the drying cycles. Part III is completely dedicated to the problem of laundry moisture estimation during drying cycles; several procedures have been tested based on available data used for simulations and, in particular, one algorithm has been chosen for the implementation on firmware and comparison with the current procedure. The Reasons for this choice will be supplied in this part as well. In the same way, Part IV will be dedicate to the other problem: laundry weight estimation during drying cycles via Soft Sensors. In this section, Functional Regression problem will be introduced explaining its value for fabric care simulation modeling problems in general. Further, in this section, the interpretability of the main proposed solutions will be analyzed and discussed to characterize them in the context of industrial modeling. In addition, chapters for results will be presented for each of the abovementioned objectives discussing differences in terms of performance and calibration effort.

Figure 1.9 shows the dissertation scheme visualizing the main contents grouped by topics. The Connection between nodes represents which method was used for a particular estimation problem: laundry moisture or laundry weight; for each applied method experimental results are reported and discussed in details (Part III and IV).

Most of the charts reported in this dissertation were normalized for industrial secrecy reasons, especially the charts that plot the evolutions of the drying variables over time.

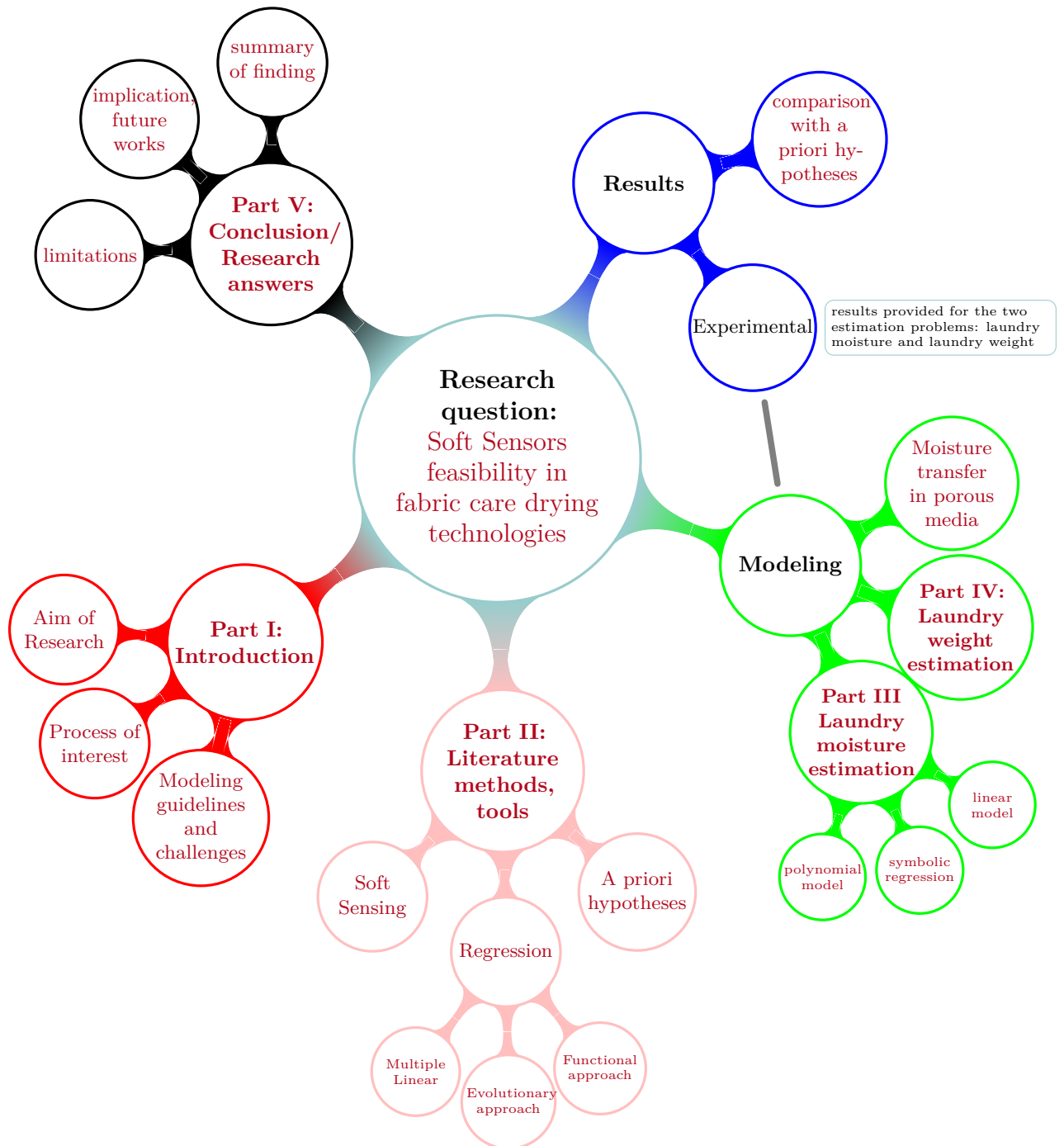


Figure 1.9: Scheme of the dissertation by topics.

Part II

Methods and tools

2

Soft Sensing

2.1 Soft Sensors in the process industry

This chapter is an introduction to Soft Sensors which, from the methodological perspective, are the main topics of this dissertation. In general terms, Soft Sensors are components of process automation and they offer an indirect or predictive way to monitor and control. They are a big part of the overall automation infrastructure (Figure 2.1) and can contribute significantly to process visibility and operator guidance. Soft sensors are based on other process observations commonly used when hardware sensors for a specific parameter are unavailable. They play the role of surrogate for hard-sensor validation when performance declines through fault accumulation.

SS exploits already in-place sensors/information, therefore representing a cost-free solution for improving product/process performances. Such technology is generally based on Machine Learning (ML) supervised techniques (Hastie, Tibshirani, Friedman, and Franklin (2005)) that exploit the availability of historical data where the relationship between inputs and output is measured.

SS technologies are used for several purposes and a relevant reference for the development of data-driven SS for the process industry can be found in Kadlec, Gabrys, and Strandt (2009) in which an introduction to the most popular SS modelling techniques as well as a discussion of some open issues in the SS development and maintenance is provided.

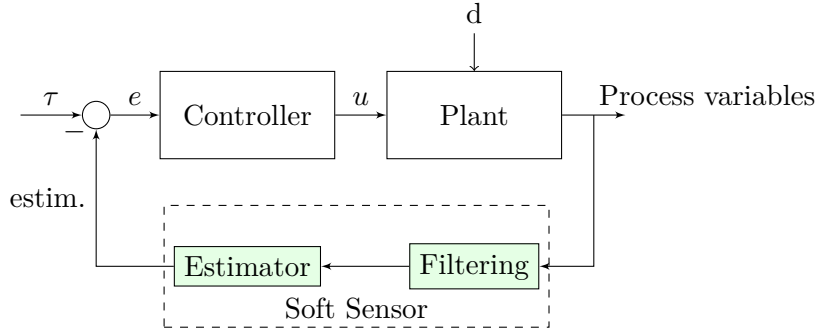


Figure 2.1: Soft Sensor in a control scheme.

This has been assumed as the main reference for the introduction of the topic in this dissertation. Another essential reference can be found in Soares (2015) which presents some ensemble learning methodologies for SS development in Industrial processes.

A large number of sensors are usually exploited by Industrial processing plants and the primary purpose of the sensors is to deliver data for process monitoring and control. Approximately two decades ago researchers started to make use of the large amounts of data being measured and stored in the process industry by building predictive models. These predictive models are called Soft Sensors. This term is a combination of the words “software” because the models are usually computer programs, and “sensors”, because the models are delivering similar information as to their hardware counterparts. The formal definition of Soft Sensors that will be considered in this dissertation is the one summarized above:

Definition 2.1.1. Soft Sensors are statistical model-based technologies used in industrial environments that provide an estimate of quantities that may be unmeasurable or costly/time-consuming to measure. They exploit signals and sensors already available from the process of interest.

Other common terms for predictive sensors in the process industry are inferential sensors (see e.g. Jordaan, Kordon, Chiang, and Smits (2004)) and observer-based sensors (Goodwin (2000)). As stated in Kadlec et al. (2009), at a very general level one can distinguish two different classes of Soft Sensors, namely model-driven and data-driven. The model-driven family of Soft Sensors is most commonly based on First Principle Models (FPM) but model-driven Soft Sensors based on extended Kalman filters (Welch and Bishop (1995)) or adaptive observers (Bastin (2013)) have also been published. First Principle Models describe the physical and chemical background of the process. These models are developed primarily for the planning and design of the processing plants, and therefore usually focus on the description of the ideal steady-states of the processes. This

limitation makes it difficult to base Soft Sensors on them. As a solution, data-driven Soft Sensors have grown in popularity in the process industry. Because data-driven models are based on the data measured within the processing plants, and thus describe the real process conditions, they are, compared to model-driven Soft Sensors, more reality related and describe the true conditions of the process in a better way. Nevertheless, there are a lot of different issues that have to be dealt with while developing data-driven Soft Sensors. These issues will be discussed later on in this thesis.

The most popular modelling techniques applied to data-driven Soft Sensors are Principle Component Analysis (PCA) (Jolliffe (2002)) in a combination with a regression model, Artificial Neural Networks (Bishop (1995)), Neuro-Fuzzy Systems (Jang, Sun, and Mizutani (1997)) and Support Vector Machines (Vapnik and Vapnik (1998)). The range of tasks fulfilled by Soft Sensors is broad. The original and still most dominant application area of Soft Sensors is the prediction of process variables which can be determined either at low sampling rates or through off-line analysis only. Because these variables are often related to the process output quality, they are very important for the process control and management. For these reasons, it is of great interest to deliver additional information about these variables at a higher sampling rate and/or at lower financial burden, which is exactly the role of Soft Sensors. The modelling methods applied to these kinds of applications are either statistical or soft computing supervised learning approaches. This Soft Sensor application field is further referred to as on-line prediction.

Other important application fields of Soft Sensors are those of process monitoring and process fault detection. These tasks refer to the detection of the state of the process and in the case of a deviation from the normal conditions to the identification of the deviation source. Traditionally, the process state is monitored by process operators in the control rooms of the processing plants. The observation and interpretation of the process state are often based on univariate statistics and it is up to the experience of the process operator to put the particular variables into relations and to make decisions about the process state. The role of process monitoring Soft Sensors is, based on the historical data, to build multivariate features that are relevant for the description of the process state. By presenting the predicted process state or the multivariate features the Soft Sensor can support the process operators and allow them to make faster, better and more objective decisions.

The typical steps of the common practice of Soft Sensor development are summarized in Figure 2.2 and reported here revising the statements of Kadlec et al. (2009). The presented procedure is rather general and refers to the case of the data which are produced in the process industry; nevertheless, it is exactly the set of procedures of interest for

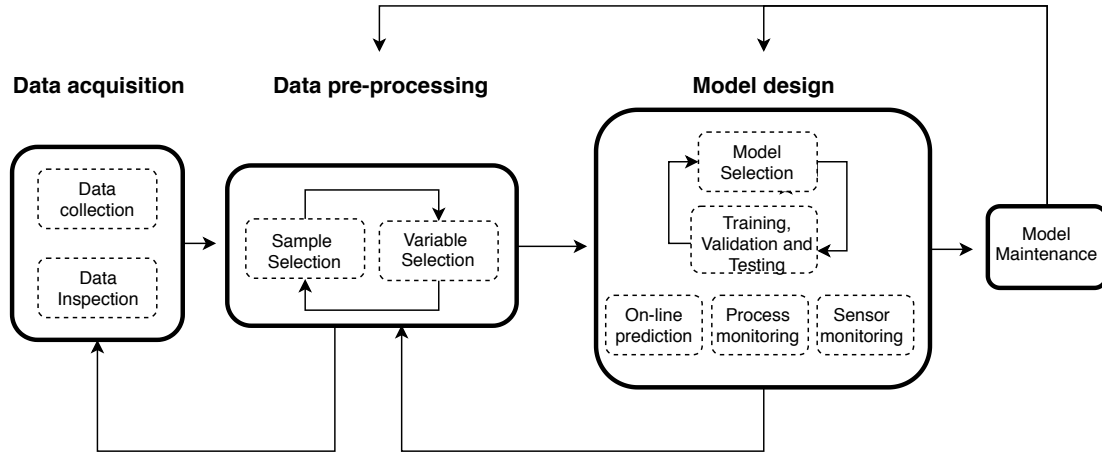


Figure 2.2: A general methodology for Soft Sensor development.

the applied research at hand here. Referring to the Figure 2.2, the data pre-processing is usually done in an iterative way, e.g. after the standardization and missing value treatment which are usually performed only once, outlier removal and feature selection are repeatedly applied until the model developer considers the data as being ready to be used for the training and valuation of the actual model. It is emphasized here that the importance of the pre-processing is critical. At the moment, the pre-processing of the data is the step that requires a large amount of manual work and expert knowledge about the process. One of the goals of my research is the improvement of this phase of SS development making it less costly in terms of time effort and without sacrificing the model interpretability as will be discussed.

The choice of the input and output variables is a crucial stage in soft-sensor development. The presence of irrelevant input variables may lead to noise, which will cause the soft-sensor model to deteriorate, and produce over-parameterization. Moreover, the higher the input variable dimension is, the heavier the model is computationally, which is not desirable. Process understanding is useful in the first screening of the potential input variables and, as for the mathematical techniques, the multivariate, clustering and cross-validation methods are often used in recognizing the relevant variables for the soft-sensor model. Sample and variable selection steps are closely connected to each other and it is often necessary to reconsider the initially selected set of samples that occur when relevant input variables are searched for. If the sample and variable selection does not satisfy the requirements set by soft-sensor developers, it is necessary to go back to the data acquisition step as shown in Figure 2.2. In the same figure the reader can appreciate how the actual soft-sensor model design consists of: (i) model structure selection, (ii) model

training, (iii) model validation and (iv) model testing. First, the selection of the model structure depends on the task at hand and the type of output desired. For instance, it can be supervised (outputs, e.g. estimates of a predefined variable) or unsupervised (e.g. recognition of data patterns corresponding to operational states). Depending on the type of data, the researcher also needs to consider whether a linear or non-linear model is required. For the case studies of this thesis, linear structures are often preferred mainly for two reasons:

- (a) as a natural first step, simple models are chosen as SSs to avoid unnecessary computational effort for the firmware. In other words, we try to use models that are as simple as possible to achieve our goals.
- (b) the interpretability of the desired SSs is one of the main requirements for our research because we need the knowledge of exactly how each predictor is related to the output in order to perform the debug and the continuous improvement of the procedures.

It is pointed out that the point (a) discussed above appears to be obvious but it has also a profound motivation related to the model overfitting problem which will be introduced in 4.1 and discussed in the following parts of the thesis. Basically, the profound motivation is that the use of good representations of the input information (which will be called features in the following) allows simple models to beat complex models and this is a fundamental guideline of the work at hand here for the design of our SSs. After the selection of the candidate models, they are trained using the pre-processed data as the model inputs. Next, the performance of the models needs to be evaluated with an independent validation data set or by means of cross-validation. After that, a model with suitable parameters defined in the model validation step is trained using both the original training and validation data and the re-trained model is tested with an unseen testing data set. The results of the model testing are analyzed, for example, using adequate statistics and graphical analysis. If model testing does not fulfill the set requirements, it is necessary to return to sample and variable selection and, thus, to try to find more representative input data set for the model. Due to the changing environment, an implemented soft-sensor model tends to gradually deteriorate and, therefore, the model testing before the implementation is not the last step of a successful soft-sensor design. Because of the importance of these steps in the SSs development especially for industrial applications, some other guidelines have been summarized below from Kadlec et al. (2009) for the main phases represented in Figure 2.2. This sort of list constitutes a systematic way for the design of SSs for industrial applications.

- Data acquisition phase: a first inspection of the available data should be made here

to identify any obvious problems (e.g. locked variables having constant value). It is also possible to verify the requirements for the model complexity and if there is enough variation in the output variable and if this can be modelled at all.

- **Data pre-processing:** this step deals with the selection of data to be used for the training and evaluation of the model. In the case of the data which are produced in the process industry, such as all our case studies in this dissertation, it's even more important the handling of missing data, outliers and selection of relevant variables (i.e. feature selection). Outliers removal and feature selection are repeatedly applied until the model developer considers the data as being ready to be used for the training and evaluation of the actual model. As already said, the importance of the pre-processing is critical in general and, in particular, for the research discussed in this thesis, basing also on my personal experience, this phase could take the 70% – 80% of the total amount of time available for the project.
- **Model design phase:** at this step the developer has to decide which kind of performance evaluation is appropriate for the problem at hand. Dealing with numerical evaluations, the most popular is the Mean Squared Error (MSE), which measures the average squared distance between the predicted and the correct value. In this research, the Root Mean Squared Error (RMSE): $\sqrt{\frac{1}{n} \sum_{t=1}^n [y(t) - \hat{y}(t)]^2}$. (where, $y(t)$ is the output sample we want to predict and $\hat{y}(t)$ is the prediction) is always preferred only because it allows the use of the same unit of measure employed for the output of interest. Visual representations are preferred as well when they provide useful information about prediction residuals, i.e., the difference between the target output and predictions. A disadvantage of visual methods is that the final decision on the adequacy of the model performance, is in any case, up to the subjective judgment of the model developer.
- **Model Maintenance:** the designed SS has to be maintained and tuned once available. The maintenance is necessary due to the drifts and other changes of the data which cause the performance of the Soft Sensor to deteriorate and have to be compensated for by adapting or re-developing the model. We point out that there is often no objective measure for assessing the Soft Sensor quality level and the judgment on whether a model works well or not is dependent on the model operator's subjective perception based on visual interpretation of the deviation between the correct target value and its prediction. In this thesis, SSs were designed using real data (after the onerous pre-processing phase) representative of the usual variability of standard drying cycles. Available data (drying tests) were prepared and selected working

closely with the laboratory experts in order to limit the performance deterioration due to data drifts and poor generalization capacity of the designed model.

Recognized problem	percentage
time-consuming maintenance	27%
need for calibration	21%
aged deterioration	15%
insufficient accuracy	13%
long dead-time, slow dynamics	10%
large noise	8%
low reproducibility	2%
others	4%

Table 2.1: Problems with hardware sensors.

Purpose	percentage
stabilize quality	37%
reduce feed and utility	28%
improve reliability with online analyzer	18%
stabilize operation	9%
reduce manual analysis	6%
avoid installing special analyzer	1%
analyze key factors	1%

Table 2.2: Purposes in using Soft Sensors.

At the end of this section, some remarks on the motivations of the use of soft sensors when no other physical counterparts are available, are discussed here to stress the significant role of data-driven solutions for problems similar to the ones addressed in this research. Even when hardware sensors can be used, some problems can be found: as an example, in Table 2.1 some of the common problems were identified as a result of a questionnaire to 26 companies in Japan (survey [jsp \(2004\)](#)) and discussed in [Kano and Fujiwara \(2012\)](#). According to this survey, maintenance and calibration are the main problems using physical or hardware sensors when it's possible to use them to measure the quantity of interest. Hence, it may be worth to consider the proposal of SSs as solutions for these scenarios. Besides, SSs are used nowadays for several different purposes and an example is shown in Table 2.2 which reports an investigation made by [Ookita \(2006\)](#). A part of this study illustrates the present state of SSs technologies in Mitsubishi Chemical Corporation and as the reader can see, the major purposes of using SSs are the

stabilization of product quality via on-line estimation as well as the reduction of energy and material consumption. Actually, this thesis matches this purpose partially, because the use of SSs for drying technology applications is aimed to improve the performance of products but it also permits a significant reduction in terms of costs, materials, and energy whenever soft sensing turns out to be effective compared to the current solutions if available. Indeed, the ultimate goal of this thesis is to provide new solutions for cycle handling to obtain efficient drying tests both in terms of fabric care performance and energy savings.

Virtual sensing technology has been successfully applied to various processes in various industries and a great number of researches has been conducted to improve the estimation performance and solve practical problems (Kano and Fujiwara (2012)). The above-mentioned studies introduce the importance and the role of SSs development in the process industry which is exactly the main topic of this dissertation. In this context, the thesis is concretely focused on the design of Soft Sensors for specific applications in the industry (household major appliances), and the purpose here is mainly the improvement of the current drying performance via the estimation of the quantities of interest which are inaccessible for technology or cost reasons.

2.2 Model-driven and Data-driven Soft Sensors

Generally speaking, it is possible to distinguish between two types of SSs: model-driven and data-driven. The former has full phenomenological knowledge about the process background and they are also called white-box models. On the other hand, data-driven models (or black box models) are based on empirical observations of the process. It is also possible to have combinations of these two major types: a typical example of such a combination is a model-driven Soft Sensor making use of the data-driven method for the modelling of fractions which cannot be modelled easily in terms of phenomenological models; these models are called grey-box.

Model-driven models are primarily developed for the purpose of planning and development of the process plants and they are based on equations describing the chemical and physical principles underlying the process. The drawback of this type of model is that their development requires a lot of process expert knowledge that is not always available. For example, for washing/drying chaotic processes where there are several combinations of laundry composition, quantities, and variability, there is often not enough phenomenological knowledge for an accurate description of the processes at hand. Another problem is that these kind of models often describe a simplified theoretical background

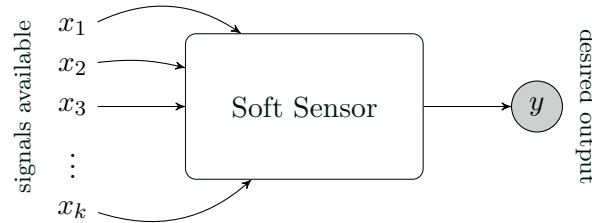


Figure 2.3: Soft Sensor as a black-box model.

of the process rather than the real-life conditions. Examples of model-driven Soft Sensors are the ones in which Kalman filters are applied (e.g. Prasad, Schley, Russo, and Bequette (2002); Doraiswami and Cheded (2014)); in these cases, a copy of the mathematical model of the plant (if available) is driven by the residual, which is the error between the measured output of the plant and its estimate generated by the Kalman filter. The Kalman gain is used as an effective design parameter to handle the uncertainty associated with the model of the physical system. Data-Driven models for Soft Sensing are used when the model of the plant is not available or is not accessible or too difficult to obtain and so the measure of interest is inferred using available data. They can be considered nowadays as an attractive modeling approach available for plant operators and embedded in automated control systems; they are based on the real-life measurements which are recorded, stored and provided as historical data. Figure 2.3 simply illustrates the information flow through a SS seen as a black-box (or grey-box) module.

In general, it must be taken into account that, the success of the deployment of a data-driven Soft Sensor goes beyond the application or design of Machine Learning schemes. The development involves different types of Data Analytics paradigms such as:

- Signal processing techniques: to prevent acquisition inconsistencies due to the precision of the equipment;
- Dimensional reduction techniques: to reduce the inputs or independent variables with less importance for the process or transform this set into another set of variables;
- Detection of novelty or outliers: decide whether an observation belongs to the same distribution or should be considered as different. Often, this ability is used to clean real data sets and it was essential in the preliminary phase of this research to obtain a reliable dataset;
- Machine learning schemes: best techniques to control overfitting, e.g. ensemble techniques.

In particular, the last point is the one which is of greater importance for this dissertation taking into account that the objective in this work is the proposal of new procedures oriented to the automatic end of the cycle and cycle handling for drying technology applications. The next section will introduce the two main estimation problems of interest and for which data-driven SSs have been studied as proposed solutions.

2.3 Soft Sensors for fabric care applications

As stated, the dissertation is focused on Data-Driven SSs for fabric care appliances; developing such models for laundry household equipment is a challenging task for many reasons, some of them have already been introduced before:

- i laboratory data, where the reference output is accurately measured and laundry characteristics are precisely described, are costly;
- ii computational capacity in current appliances is limited, posing a hard constraint on the solutions to be embedded in the equipment.

In previous works and also for the SSs designed for this dissertation, we overcome these issues by exploiting laboratory data already collected for other product development purposes and by exploiting regularization ([Scholkopf and Smola \(2001\)](#)), a Machine Learning framework that will be described in details in the following parts and which allows, in some cases, to provide effective sparse linear models that are easily implementable.

Target measure	unit	WM	WD/TD	offline	online
Energy consumption	[J]	✓		✓	
Water consumption	[l]	✓		✓	
Virtual Energy Label	[J]	✓	✓	✓	
Laundry dry weight	[Kg]	✓	✓		✓
Laundry Load typology	[-]	✓	✓	✓	✓
Water Extraction	[l]	✓	✓	✓	
Laundry moisture	[%]		✓		✓

Table 2.3: Summary of past and current fabric care R&D projects based on SSs development; WM=Washing Machines, TD=Tumble Dryers, WD=Washer/Dryers. Two rows are highlighted because they refer to the SSs of interest in this dissertation.

In order to summarize the set of SSs developed so far by fabric care R&D (already implemented or on-going projects), Table 2.3 shows which are the features for each case study. As represented, SSs have been developed not only for drying technologies applications which are the main ones of interest for this thesis but also for washing cycles

in washing machines as well. As a matter of fact, the methodologies used to create SSs are similar for washing/drying machines as regards signal processing and data analysis, therefore, every time a method turns to be useful for a specific goal on a specific machine, sometimes it can be adapted to work for different technology modifying only a few parts of the entire SSs development procedure. Table 2.3 specifies which is the target output for each SS (it can be a scalar value or a time series - not specified here because it's the object of the next chapter), it's also specified whether the SS is model-driven or data-driven and this special feature is fundamental to interpret the challenges mentioned before; in fact, when an online implementation is required, the constraints about computational burden and limited memory becomes of primary importance and the set of possible SSs solutions shrinks.

The target measure specified in Table 2.3 describes the output of the model provided by a specific SS designed for that goal, e.g. to compute/predict energy consumption in Washer Machines and/or in Tumble Dryers, Washer-Dryers. The Model developed to predict water consumption is specifically thought for Washer Machines as well as models for water extraction (during centrifuge). Virtual Energy Label is an on-going project oriented to the prediction of consumptions for the computation of parameters used by institutes to validate the energy label associated with the machine exploiting several models and a huge amount of data from the factory. SSs have been designed also for the online prediction of laundry load typology using data from WD and WM, while the two highlighted rows in Table 2.3 show the two outputs of interest in this dissertation and described in details in the following sections.

As will be explained later, all the SSs summarized in the table can be considered grey-box models because the expert knowledge of the process is inherent in the feature engineering phase (which is done manually) where the information inside inputs signals is conveniently represented and summarized. For the two cases highlighted above, the focus of this thesis is the use of several methods to bypass the hand-craft feature engineering phase which is always time-consuming, exploiting input signals as they are in order to make the SS development leaner and less subjective.

3

Algorithms for drying technology applications

3.1 Automatic End-of-Cycle determination

In this chapter, an introduction to the main algorithms currently used for the development of drying technology applications in fabric care is presented to the reader. The comprehension of the explained procedures here is central to appreciate the novelty of this dissertation which is focused on the proposal of new methods to solve these problems. One of the most important controls implemented on Tumble-Dryer (TD) and WD-HP is the automatic End-of-the-Cycle (EoC) procedure. It is essential to ensure the required and claimed drying performance; the main features of this practice are summarized below. A Concept that will occur often in this part of the dissertation is the one introduced by the following definition.

Definition 3.1.1. Automatic End-of-the-Cycle (EoC) goal: being more accurate in the automatic determination of the cycle-end which correspond to the instant in which the laundry is perfectly dried (when the estimated moisture reaches the predefined moisture target¹, the cycle terminates). **An automatic dry algorithm aims to infer when the target final humidity is reached and stop the drying cycle.**

A higher precision for the automatic dry algorithm means:

¹For example, 0% is the moisture target for Cotton fabric.

- more reliable results of the Energy Label (an image is provided in Figure 3.1) compliance tests;
- better evaluation of the products by the consumer magazines.

For these reason the precise automatic EoC determination is one of the fundamental purposes of the research described in this dissertation. An explanation of the current

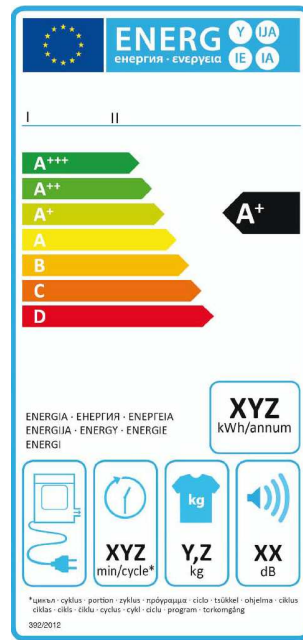


Figure 3.1: The Energy Label. Image taken from EL (2012).

procedure adopted on machines to perform this task is provided here and will be discussed again in the thesis to compare new proposed solution for the EoC goal. The current EoC algorithm adopted online on drying machines (TD - WD) is based on monitoring of a signal which is elaborated exactly for this purpose. When such a signal reaches a predefined set of thresholds, the drying cycle end is decreed and a final cooling phase starts in order to avoid the overheating of the modules of the machine. Thresholds for the EoC have to be set offline by functionality experts during calibration phases of the drying algorithm. A set of thresholds is set for the main drying programs and machines of interest, e.g. machines with different maximum drum speed or motor. Considering all these variants, the calibration procedure is dependent on the quantity and quality of drying tests available at the time when this phase takes place. In fact, if few tests (less than 10) are available to perform calibration, additional acquisitions have to be planned with a significant time engagement. Indeed, complete drying cycles (to get the laundry moisture to the target) are necessary for calibration and each one is time-consuming for

laboratory facilities and resources.

The signal used to set the end of the automatic drying cycle in a Washer-Dryer (WD) will be called the EoC signal and is simply described as a function of the following input signals:

1. the air temperatures at the drum inlet and outlet (acquired by NTC sensors);
2. the motor torque.

The resulting signal is calculated by mainboard firmware every 1 sec and it is simply the sum of mechanical and thermal components² available on machines.

Figure 3.2 depicts how the EoC signal is computed: Figure 3.2a shows the thermal components, i.e. temperature signals from 3 different NTC sensors (details omitted) while Figure 3.2b illustrates the mechanical components, i.e. the elaborated signal from motor torque. In detail, the 3 temperatures are combined together properly to form the EoC first component which is called EoC part1 while the mechanical component is processed to obtain the signal EoC part2. In the end, the two components are combined to obtain the EoC signal depicted in Figure 3.3. When the EoC signal reaches the represented threshold, the cycle end is decreed. A different type of visualization is provided in Figure 3.4 where the EoC signal is represented for different laundry load sizes, i.e. each drying test has a different nominal load³, for example 1[kg] of laundry etc. and for each case the EoC is illustrated. As an observation, the reader can notice how the range of values for the EoC signal differs when the load is small (C) or medium-large (A, D and B). From the observation of Figure 3.4 2 remarks are pointed out here:

- (i) the variation of the signal in relation to the laundry load size: the shape for small loads (like load A) is different from the one shown for medium or large sizes (B-D);
- (ii) in general the EoC signal varies more approaching the real cycle end (when the moisture target is reached) and this is exactly the behavior that is desired for the EoC purpose, e.g. to obtain a signal which is informative about the cycle end.

The moisture reference represented on the x-axis in Figure 3.4 has been made possible considering elaborations from the rough laundry weight signal which is acquired in the laboratory during drying using a balance under the machine. Clearly, such a signal is not available online and is made available from laboratories in order to evaluate the

²*omissis*. Details will be omitted here because of intellectual property rights.

³From now on, the term nominal load will be used to mean the weight of the dry laundry, i.e. laundry at the moisture target.

EoC performance measuring at which moisture the algorithm stops the cycle and the distance to the moisture target fixed by the selected program. Other details involving the laundry moisture reference reconstruction from the weight evolution will be provided in 5.2. A significant amount of time was spent on the improvement and the setting of a suitable moisture reference from rough weight signal solving problems related to corrupted acquisitions and noise. For example, in order to consider only the amount of water inside the laundry without the amount circulating in the machine pushed by the drain pump (Figure 1.3a), only a subset of the base weight signal must be taken into account, filtered and interpolated; a systematic procedure was designed to handle several scenarios on laboratory tests (other details follow in 5.2).

Dealing with details about the EoC determination on WD, to select properly the EoC signal threshold online it is necessary to identify the laundry load quantity and the machine temperature which is indicator for the status of the machine (if there was another drying cycle before the current one; in this case the drying test is said to be “hot” because it takes place on a heated machine).

In particular, at the beginning of the drying cycle, just before a certain motor impulse, the inlet temperature inside the drum is read and compared to a temperature predefined threshold to discriminate 2 machine temperatures (Cold, Hot). Then, a measure of inertia due to the laundry inside the drum is calculated (the computation is actually repeated more times after drum movements and then a mean measure is considered) and compared to some Inertia levels to discriminate laundry quantity levels (Low, Medium, High, ExtraHigh).

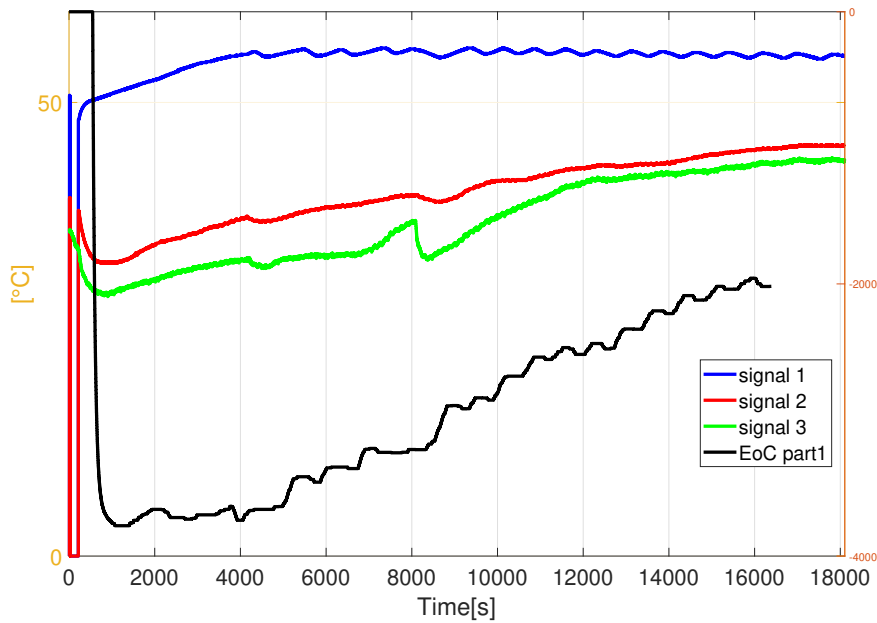
Examples of inertia values for different drying cycles in a washer-dryer machine are represented in Figure 3.5 where cycles are grouped by nominal laundry load (from left to right: small(A), medium(B,C) and large(D))⁴.

In the case study analyzed here, i.e. a washer-dryer machine, a dedicated phase at the beginning of the cycle is scheduled to make the inertia measures (more than once in order to obtain a reliable value). According to the currently implemented algorithm on washer-dryer machines, the laundry load weight information is provided by a procedure which includes an initial impulse to the drum to adhere to the laundry and then some other movements to obtain the inertia due to the laundry. The laundry dry weight estimation problem is crucial for the majority of algorithms designed for cycle handling in fabric care R&D projects, for both drying and washing technology. As a matter of fact, in general, the washing/drying performance together with the energy-saving is influenced by the amount of laundry load inside the drum during the cycle and so, its precise knowledge

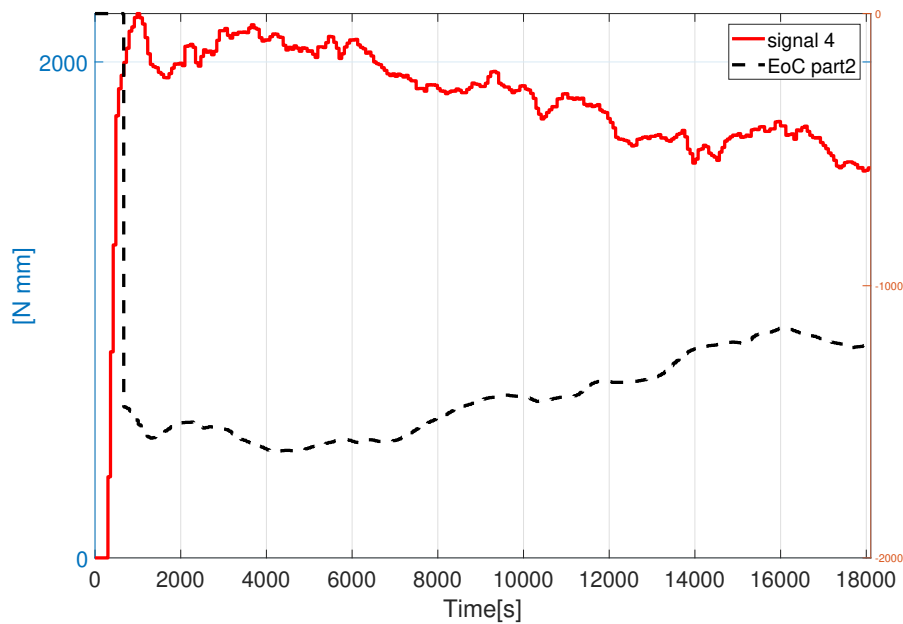
⁴omissis: other details about the computation of inertia levels are omitted here because of intellectual property rights.

is mandatory for a suitable treatment of the process. Given the importance of this topic, a dedicated Section has been prepared in this dissertation (Section 3.3). To sum-up, the EoC signal is a function of: time elapsed, load quantity, machine temperature and the particular selected program (on which EoC thresholds depend). For what has been said, beyond end-of-cycle performances, the main problem related to the current algorithm is due to the time spent by the operator who sets the thresholds for each case, so the current procedure depends on a time-consuming calibration process and one purpose of the new proposals is to avoid/improve the calibration time effort associated with the drying algorithm.

As a matter of fact, the algorithm discussed so far and used nowadays is not based on a measure or estimation of the most important output reference of a drying process, i.e. the laundry moisture content during a drying cycle. New proposals are focused on the estimation of this particular signal which is difficult or costly to measure directly; for this reason, SSs have been developed as new solutions for the problem at hand.



(a) Thermal components of the EoC signal. Large laundry load size acquisition.



(b) Mechanical components of the EoC signal. Large laundry load size acquisition.

Figure 3.2: The EoC signal components.

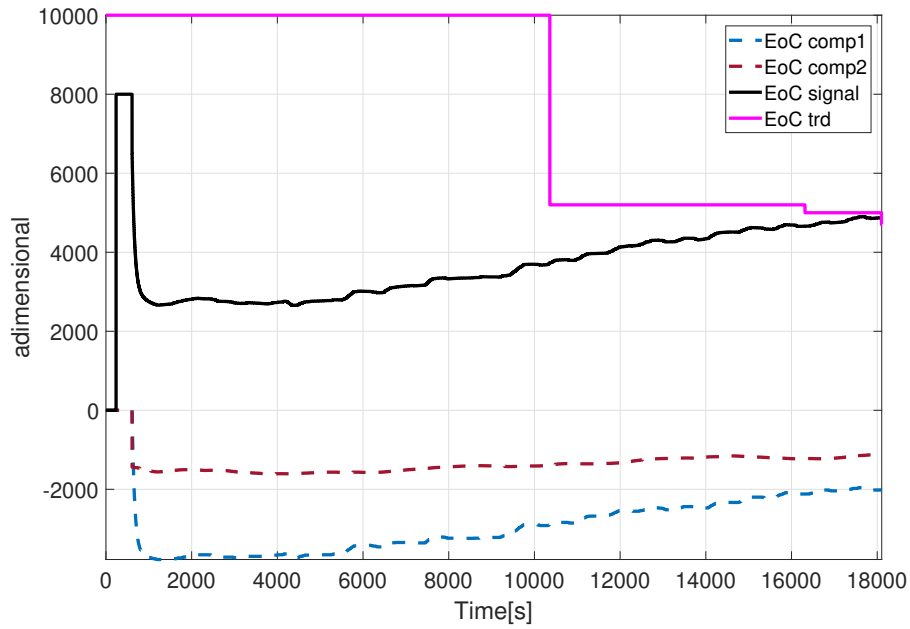


Figure 3.3: EoC signal with threshold for the cycle end detection.

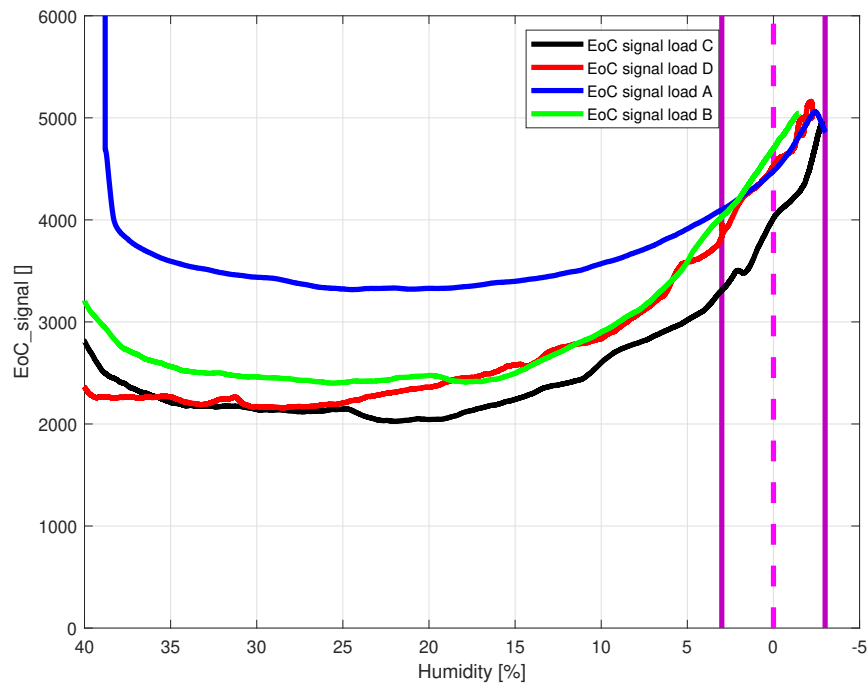


Figure 3.4: EoC signal vs laundry moisture (Humidity) obtained from elaborations of the weight signal available from laboratory tests.

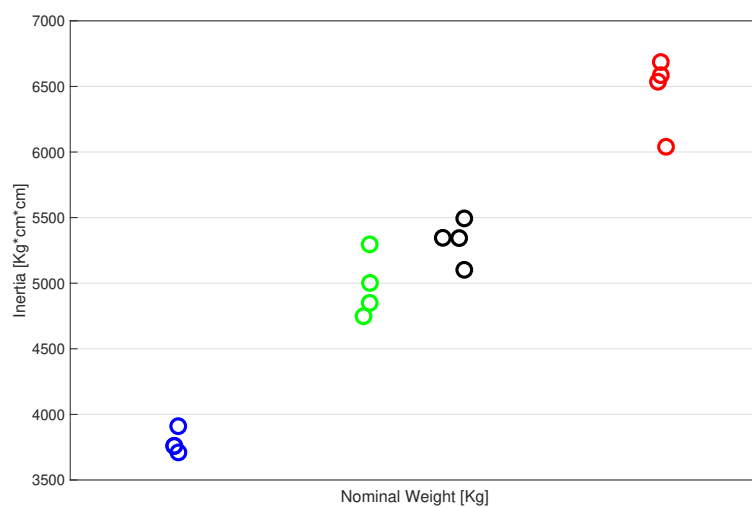


Figure 3.5: Inertia levels for different drying cycles. from left to right: small(A), medium(B,C) and large(D); nominal weights omitted.

3.2 Time-to-End algorithm

Another essential procedure implemented on machines of interest is the so-called Time-To-End (TTE) algorithm. This function is specifically designed to provide information to the user at the beginning of the (drying) cycle and then to be updated for a more precise cycle handling. Such a piece of information is clearly the residual time needed to reach the laundry moisture target. It turns out that the first computation and the update of this time amount are critical for the entire cycle control. Clearly, the two concepts of EoC and TTE are related to each other, but actually, they are two different algorithms. In order to make it absolutely clear to the reader the definition of TTE is provided:

Definition 3.2.1. TTE goal: know beforehand accurately and precisely how much time is needed to reach the moisture target using estimations available in the first part of the cycle in order to provide an indication to the user interface.

The difference between EoC and TTE is the fact that for the latter, the goal is to predict the progress of the cycle, i.e., a moisture estimation which depends on the current value of the inputs is not enough. In Figure 3.6 a frontal image of the WD-HP of interest is given. It shows the user interface in which the TTE indicator is displayed and provided to the user. It is pointed out here that the TTE algorithm implementation is not just for a visualization purpose but for an optimal cycle handling through the knowledge of the essential information for the entire process. A more vivid representation of the evolution of TTE computation during drying cycles is provided by Figure 3.7 which reports the TTE profile for different cycles distinguished by the laundry nominal load. The first computation of TTE is given at the beginning of the cycle and depends on the determination of the inertia value. In fact, the TTE evolution represented in Figure



Figure 3.6: User Interface, Electrolux WD-HP.

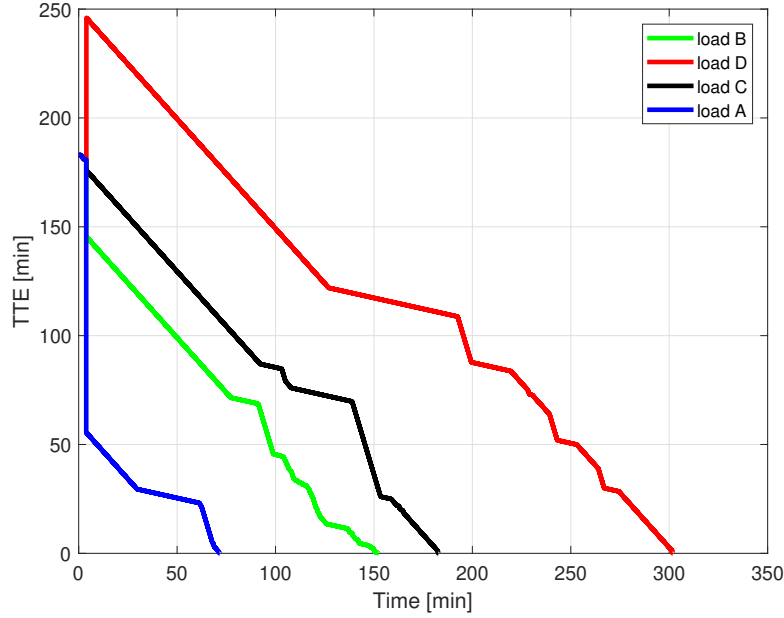


Figure 3.7: TTE representation for drying tests with different laundry load size.

3.7 is the estimate of the time duration from the last inertia measure to the start of the cool-down phase. On the x-axis the reader can appreciate the elapsed time while on the y-axis there is the TTE estimate. The first estimated value is proportional to the average value of the inertia measures computed during the first minutes; in detail, a particular dependence (omitted) is exploited and provides the first TTE estimated value from the inertia measures. Parameters are elaborated offline and written on a firmware dedicated module; they depend on the selected drying program. The rest of the TTE is updated every minute using elaborations from the EoC already mentioned signal (omitted). Figure 3.7 shows the TTE displayed to the user through the user interface of the machine; examples provided for drying tests with several laundry loads: load A, B, C and D in Figure 3.7.

3.3 Laundry weight determination

For what I said earlier, a suitable laundry weight estimation is essential for lots of R&D fabric care projects related to drying or washing technologies. As anticipated in Section 1.1, few Machine Learning (ML)-based works on SSs for household major appliances are currently available in scientific literature and one of the reasons is the difficulty in collecting observations for the design of a ML algorithm. In the case study used in this

dissertation, observations are complete drying tests which last 3 hours at least. The main rationale is that we are interested in the prediction of the behavior of an entire cycle and so we need to acquire complete tests in laboratories. Considering that the priorities of manufacturers are often focused on the development of several projects related to production issues, the preparation of a dedicated test plan for the research purpose at hand, turned out to be troublesome.

As regards the documentation available in the patent literature about general methods for the automatic estimation of the laundry load size, Smargiassi, Potena, Minni, Concettoni, and Animobono (2013) proposed a method based on the idea of estimating the weight value (e.g. the weight x of the contents of the oscillating assembly of the washing machine) by starting from the measured quantity (e.g. an electric parameter y , in particular, voltage of the electric signal generated by a magnetic sensor, such as the Hall effect sensor⁵) through a mathematical function $x = f^{-1}(y)$; some assumptions on the form of f are made according to the a priori knowledge of the relationship between the electric parameter and the estimated weight. As another example, the patent Pelczer (2012) describes an algorithm for obtaining a reading of a magnetic sensor (like Hall effect sensor) for deriving the weight of the contents of the oscillating assembly of the washing machine from said reading.

The algorithm is based on theoretical curves of the suspension system (springs and dampers), which however do not take into account the tolerances of the said system that arise once the system has been installed in a certain washing machine and cannot be foreseen a priori. This may lead to systematic errors in the detection of the weight of the contents of the oscillating assembly, since it is not possible to perform a calibration procedure capable of adapting the detection of the magnetic sensor to the different operating conditions of said magnetic sensor when installing in different washing machines (this also applies to machines of the same model).

Other patents, for example, employ a dynamic model of the basket/clothes and motor in performing clothes load estimation (Didier and Geay (1988)) and are used for determining, for example, the mass of wet laundry in a washing machine drum being drivable by an rpm-controlled universal motor, with the aid of a controller variable; in Blauert, Boldt, Wobkemeier, Pilgram, and Rehfuess (1996) for example the mass of laundry is determined during the startup phase at constant acceleration; in that phase, the phase angle and the magnitude of the supply voltage are measured, and are used to calculate the mass inertia of the laundry by taking engine constants and the acceleration performance of the empty drum into account. Chbat and Badami. (2000) use a dynamic model of moving

⁵It is a transducer that varies its output voltage in response to a magnetic field; commonly seen in industrial applications.

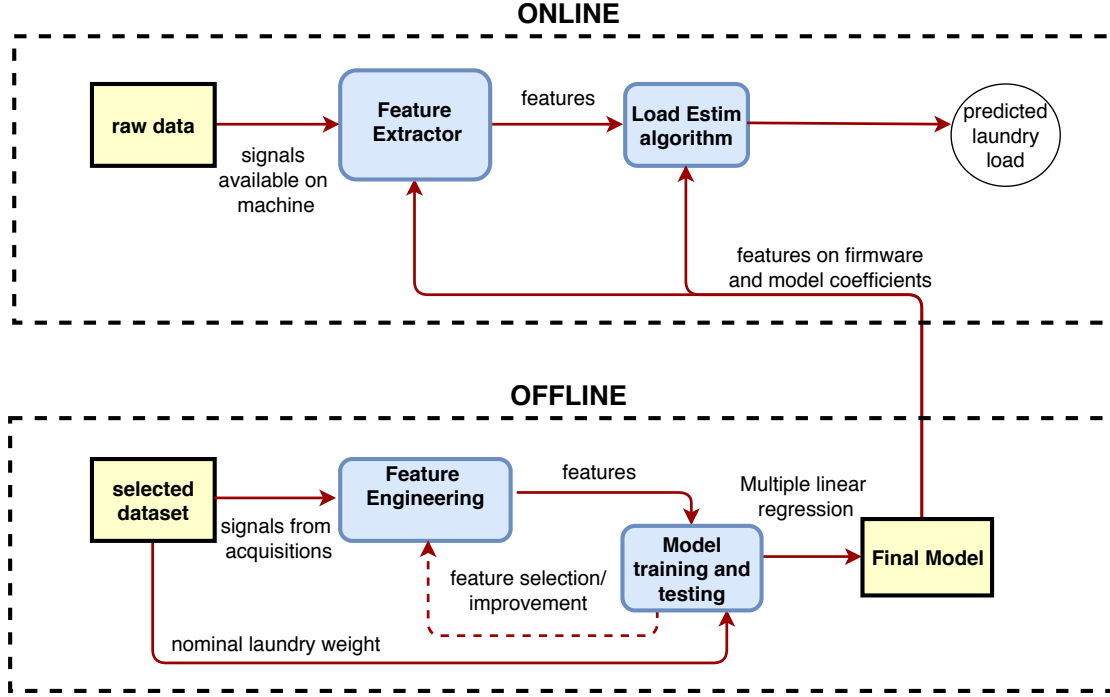


Figure 3.8: Laundry load estimation performed by the current algorithm on dryer machines.

components in the washer, the clothes load is estimated by obtaining a value of the moment of inertia of the clothes present in the washer basket, and using the inertial value, with a lookup table or an inference system, to estimate the actual clothes weight. In this method, at least two readings of the rotational speed of the basket assembly are taken by a velocity or position sensor during the acceleration phase. The readings are taken to compute the rate of acceleration of the basket assembly; the clothes load is estimated using the sensed and calculated acceleration rate and a reading of the motor phase angle sensed, by a motor phase angle sensor, during the same period of time that the basket assembly speeds were obtained. The method and apparatus require motor phase sensors and either velocity or position sensors in controlling the Washer operating condition.

All methods reported in patent literature are based or related to the knowledge of the process of interest and their implementations suffer from the adaptability to different operating conditions of washer and dryer machines. The current procedure implemented on our washer-dryer machines is based on inertia measures as shown in Figure 3.5 and explained in Section 3.1. This is not the only adopted practice, in fact, on our dryer machines and washer machines as well, soft sensors (regression-based statistical models) provide an estimation of the nominal laundry weight at the beginning of the cycle. Such

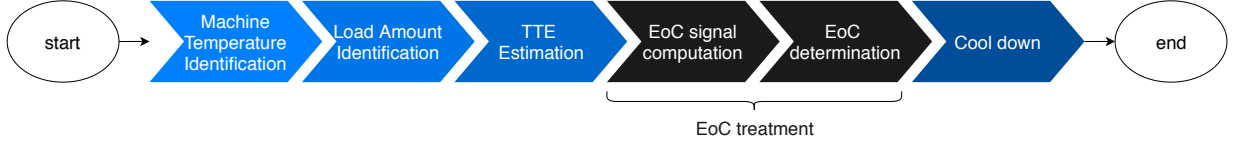


Figure 3.9: Drying cycle handling, WD-HP case study.

technology here can be exploited to obtain an estimate of the output (load weight) adaptable as much as possible to several situations (different drying cycles and change of some components of the machine) taking advantage of historical data available, avoiding additional hardware and consequent costs. The problem of laundry load estimation has been addressed in several projects implementing linear regression models on firmware in the past few years. In general, according to the current algorithms implemented for drying cycle handling in dryer machines, right after the user hits the start button, preliminary phases are launched (a comprehensive representation of the cycle handling is provided in Figure 3.9). During the Load amount identification phase, meaningful features from specific signals are extracted and combined together (today via multiple linear regression models) in order to obtain a guess of how much laundry is present within the machine. Regression models have been chosen for the specific need of the problem which is the estimation of numeric values, i.e. the laundry dry quantity in $[kg]$.

Load estimation obtained with such models have been used to update the TTE determination in the first part of the cycle both in dryer and washing machines given that the structure of the model and the phases to get it to remain the same regardless the washing or drying purpose. The load estimate can be refined with time, depending on the time at which it becomes necessary. Currently, it becomes available after the first minutes from the cycle start and is never updated anymore during drying cycle (details omitted because of intellectual property rights). A comprehensive diagram of the procedure adopted so far on dryer machines is visualized in Figure 3.8. The entire model is designed offline using data available from laboratories. Through supervised statistical learning approaches (explained in Section 4.1) and feature engineering, the information is treated to obtain the estimation goal with the minimum error (as explained in the next sections of the thesis). Each part of the algorithm will be reconsidered in part IV dedicated entirely to this problem.

Considering the difficulties and tasks described above and taking into account the variability due to several possibilities for the laundry composition inside the drum during drying cycles in machines of interest (WD or household dryers as well), the challenge of the research activities described in this thesis, is not limited to the determination of the

best model, but also implies an accurate estimation of the parameters belonging to that model using information available on household appliances using data-driven techniques exploiting real data. A visual explanation of an entire drying cycle handling is shown in Figure 3.9 where the main blocks represent the main algorithms discussed above. It is proposed here to the reader to clarify the right sequence of each procedure and how they are related to the definitions introduced before.

4

Methodologies

4.1 Multiple Linear Regression (MLR)

The research theme discussed in this dissertation is focused on the development of SSs for the determination of laundry moisture and laundry weight quantities. For this reason, this chapter is dedicated to introducing the common ML framework of techniques exploited to achieve the already mentioned estimation goals; regression models are indeed the topic of this Section taking into account the requirement which is the determination of numerical quantities of interest.

The first need in this context is defining what regression analysis is. Taking into account considerations from [Berk \(2008\)](#), a common conception in many academic disciplines and applications is the association of regression analysis with some special case of the generalized Linear model: normal (linear) regression, binomial regression, Poisson regression, or other less common forms. For any of these formulations, inferences are often made beyond the data to some larger finite population or a data generation process. Commonly these inferences are combined with statistical tests, and confidence intervals. It is also popular to overlay causal interpretations meant to convey how the response distribution would change if one or more of the predictors were independently manipulated. From another point of view, statisticians and computer scientists typically start further back. They consider Regression problems about conditional distributions. Citing [Cook](#)

and Weisberg (2009): the goal is to understand as far as possible with the available data how the conditional distribution of some response y varies across subpopulations determined by the possible values of the predictor or predictors. Regression analysis is actually defined using conditional distribution: $Y|X$, but no claims are made about how the data were generated. The regression analyses apply solely to the data at hand and are not generalized to some large set of observations. Regression analysis may be enhanced by such extensions, although they do not go to the core of how regression analysis is defined. In practice, a richer story would likely be obtained were additional predictors introduced but that too is not a formal requirement of regression analysis. There is nothing in regression analysis that requires statistical inference: inferences beyond the data on hand, formal tests of null hypotheses, or confidence intervals. And when statistical inference is employed, its validity will depend fundamentally on how the data were generated. If there is to be a cause-and-effect overlay, that too is a subject-matter or policy call unless one has conducted an experiment. When the data result from an experiment, the causal variables are determined by the research design. Regression analysis can serve a variety of purposes.

1. first, the goal is solely a description of the data at hand. In the first level, Regression could be effectively assumption-free. Too often, a description is undervalued as a data analysis tool perhaps because it does not employ much of the apparatus of conventional statistics. According to Berk and Freedman (2003), the view taken here is that p-values and all other products of statistical inference can certainly be useful, but are worse than useless when a credible rationale cannot be provided.
2. For a second level regression analysis, statistical inference is the defining activity. Estimation is undertaken using the results from the previous level, often together with statistical tests and confidence intervals. Statistical inference forms the core of conventional statistics, but proper use with real data can be very challenging; real data may not correspond well to what the inferential tools require. For the statistical procedures emphasized here, statistical inference will often be over-matched. There can be a substantial disconnect between the requirements of proper statistical inference and adaptive statistical procedures such as those central to statistical learning.

A focus on conditional distributions will be a central feature in all that follows. One does not require generative models, statistical inference, or causal inference. On the one hand, a concentration on conditional distribution may seem to limit. On the other hand, a concentration on conditional distributions may seem liberating. As reported in Berk

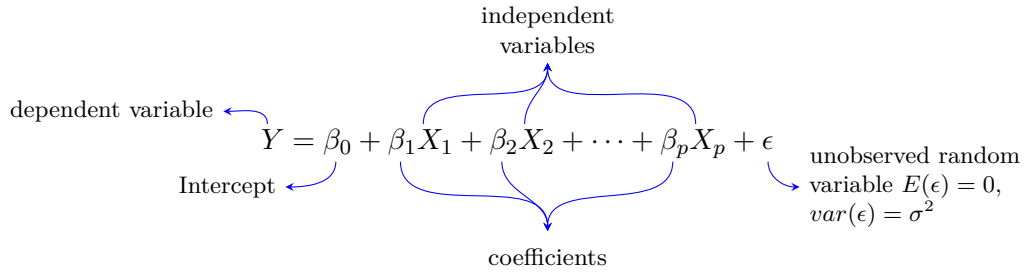


Figure 4.1: Multiple Linear Regression (Soft Sensor).

(2008), in practice, both can be true and be driven substantially by the limitations of conventional modeling.

The starting point to describe the methodologies of interest in this dissertation is the description of multiple linear regression. As stated in Harrell Jr (2015), the ordinary multiple linear regression model is frequently used and has parameters that are easily interpreted; exactly for this reason, this method meets the needs of our projects. In this part we study a general class of regression models, those stated in terms of a weighted sum of a set of independent or predictor variables. It is shown that after linearizing the model with respect to the predictor variables, the parameters in such regression models are also readily interpreted. Also, all the designs used in ordinary linear regression can be used in this general setting as seen in Harrell Jr (2015). Besides describing and interpreting general regression models, we describe how the three types of assumptions of regression models can be examined. First, we introduce notation for regression models. Let Y denote the response (dependent) variable, and let $X = X_1, X_2, \dots, X_p$ denote a list or vector of predictor variables (also called co-variables or independent, descriptor, or concomitant variables). These predictor variables are assumed to be constants for a given individual or subject from the population of interest. Let $\beta = \beta_0, \beta_1, \dots, \beta_p$ denote the list of regression coefficients (parameters). β_0 is an optional intercept parameter, and $\beta_1, \dots, \beta_p$ are weights or regression coefficients corresponding to $X = X_1, X_2, \dots, X_p$. We use matrix or vector notation to describe a weighted sum as reported in Figure 4.1 where is implied $X_0 = 1$.

It is standard to assume that the errors are normally and independently distributed with zero expectation and constant variance σ^2 . If the $X_i, i = 1, \dots, p$ are sampled along with the Y_i , rather than fixed by design as in an experiment, then it is additionally assumed that X_i are distributed independently of the ϵ and this last assumption may be constructed structurally. In fact, we require that omitted causes of y , which are thus components of the error, that are not themselves effects of the X_i , are linearly

uncorrelated with the X_i . This last assumption cannot be checked from data because the least-squares fit ensures that the residuals, which estimate the errors, are uncorrelated with the X_i in the sample.

A regression model is stated in terms of a connection between the predictors X and the response Y . Let $C(Y|X)$ denote a property of the distribution of Y given X (as a function of X). For example, $C(Y|X)$ could be $E(Y|X)$, the expected value or average of Y given X . We can now define a regression function as a function that describes interesting properties of Y that may vary across individuals in the population. X describes the list of factors determining these properties. Stated mathematically, a general regression model is given by $C(Y|X) = g(X)$. We restrict our attention to models that, after a certain transformation, are linear in the unknown parameters, that is, models that involve X only through a weighted sum of all the X s.

The general linear regression model is given by $C(Y|X) = g(X\beta)$. For example, the ordinary linear regression model is $C(Y|X) = E(Y|X) = X\beta$, and given X , Y has a normal distribution with mean $X\beta$ and constant variance σ^2 . It is important to note that the general linear regression model has two major components: $C(Y|X)$ and $X\beta$. The first part has to do with a property or transformation of Y . The second, $X\beta$, is the linear regression or linear predictor part. The method of least squares can sometimes be used to fit the model if $C(Y|X) = E(Y|X)$. Other cases must be handled using other methods such as maximum likelihood estimation or nonlinear least squares.

Linear least-squares regression analysis, in general, is susceptible to a variety of difficulties and makes often unreasonable assumptions about the structure of the data. For this reason, regression diagnostics are techniques available for exploring problems that compromise a regression analysis. Here I'll report an introduction to such techniques sending the reader to the following references: [Fox \(1991\)](#) and [Bollinger \(1981\)](#). In particular, the following selection of diagnostics is presented as a reference for what will be reported in the prosecution of the dissertation:

- (a) Collinearity between independent variables;
- (b) Non-normality of residuals;
- (c) Nonlinearity of the model;
- (d) Non-constant error variance.

Dealing with the problem (a), and taking [Yan and Su \(2009\)](#) as a reference, the collinearity refers to the situation in which two or more independent variables in a multiple linear regression model are highly correlated. Referring to the already used notation for the

multiple linear regression, the collinearity occurs if the independent variable is highly linearly correlated to another one or more independent variables. In other words, X_i can be almost linearly expressed by one or more other column vectors in X . In this situation, the matrix $X'X$ is ill-conditioned or near singular. If there is perfect collinearity among column vectors of X then the matrix $X'X$ is not invertible. When the column vectors of the design matrix X are highly correlated, then the matrix $X'X$ becomes ill-conditioned and the least-squares estimator become less reliable. Thus, the collinearity in column vectors of X will have a negative impact on the least-squares estimates of regression parameters and this needs to be examined carefully when doing regression modeling.

The normality assumption for the error term represented in Figure 4.1 has already been introduced. Now, assuming that the residual of the linear regression model mentioned is defined as the difference between the observed response variable Y and the fitted value $\hat{Y}$, i.e., $e = Y - \hat{Y}$. Residual is an important measurement of how close the calculated response from the fitted regression model to the observed response. The purposes of the residual analysis are to detect model misspecification and to verify model assumptions. Residuals can be used to estimate the error term in the regression model, and the empirical distribution of residuals can be utilized to check the normality assumption of the error term (Q-Q plot that will be introduced in the prosecution of the thesis), equal variance assumption, model overfitting, model under-fitting, and outlier detection. Fitting the model by least-squares we have $\beta = (X'X)^{-1}X'Y$ and writing regression residual in a vector form:

$$e = Y - X(X'X)^{-1}X'Y = (I - X(X'X)^{-1}X')Y = (I - H)Y$$

where $H = X(X'X)^{-1}X'$ is called the hat matrix and contains useful information for detecting outliers and identifying influential observations. It maps Y in $\hat{Y}$ and it is a projection matrix, i.e. it describes the influence each response value has on each fitted value. Now, $(I - H) = (I - H)^2$ because of the properties of H as a projection matrix (symmetric and idempotent), then: $Cov(e) = (I - H)Var(Y)(I - H)' = (I - H)\sigma^2$ and $Var(e_i) = (1 - h_{ii})\sigma^2$. Thus, if the sample size is much larger than the number of the model parameters, i.e., $n \gg p$, or sample size n is large enough, h_{ii} will be small as compared to 1 because H is a projection matrix (reference: [Hoaglin and Welsch \(1978\)](#)). As a result, a residual in multiple regression model behaves like an error if the sample size is large. However, it is not true for a small sample size. So, it is unreliable to check normality assumption using the residuals from a multiple regression model when the sample size is small as will be remarked in the next sections of this dissertation.

An example of tests for non-normality is the residual plot: e_i 's against the fitted values

y_i 's, which is a simple and convenient tool for regression model diagnosis. The residuals evenly distributed on both sides of $Y = 0$ imply that the assumptions $E(\epsilon) = 0$ and constant variance $Var(\epsilon_i) = \sigma^2$ are appropriate. A curvature appearance in the residual plot implies that some higher-order terms in regression model may be missing. Besides, periodical and curvature residuals may indicate that the possible regression model may be piecewise and some higher-order terms in the model may be missing. Another tool for non-normality check is the histogram of residuals in order to verify if the shape is similar to a Gaussian one. Examples of the use of these tools will be provided in this thesis taking into account that, when model assumptions are violated the model is no longer valid and statistical inference based on the model is not reliable anymore.

Dealing with nonlinearity, the assumption that $E(\epsilon)$ is everywhere zero implies that the specified regression surface captures the dependency of y on the x s and violating the assumption of linearity, therefore, implies that the model fails to capture the systematic pattern of relationship between the dependent and independent variables. The partial residual plot or the partial regression plot graphically helps depict the linear or nonlinear relationship between the response and an individual predictor after adjusting for other predictors. As regards non-constant error variance, one of the assumption of the regression model is that $Var(\epsilon) = Var(Y|X_1, \dots, X_k) = \sigma^2$ is everywhere the same. Non-constant error variance is often termed heteroscedasticity. For the sake of clarity, in classical regression, the equal variance assumption simply states that, given X , the conditional variance of the error terms is constant for all observations. In other words, whenever X varies, the corresponding response Y has the same variance around the regression line. This is the so-called homoscedasticity. Referring to the weighted least square estimation, this assumption implies that all Y values corresponding to various X 's are equally important and should be evenly weighted. In contrast, the condition of the error variance not being constant over all observations is called heteroscedasticity. In general, there are several reasons why stochastic disturbance terms in the model may be heteroscedastic:

- Response variable may change its magnitude according to the values of one or more independent variables in the model. Therefore, it may induce a heteroscedastic error.
- As data collecting techniques improve, it is likely to decrease the variation and to commit fewer errors.
- Outlier observations are much different in relation to the other observations in the sample. The inclusion or exclusion of such observations, especially if the sample

size is small, can substantially alter the results of regression analysis.

- Very often the heteroscedasticity may be due to the misspecified regression model such that some important variables are omitted from the model.

As a consequence of heteroscedasticity, the ordinary least squares (OLS) estimators and regression predictions based on them remain unbiased; regression predictions will be inefficient as well. Moreover, because of the inconsistency of the covariance matrix of the estimated regression coefficients, the tests of hypotheses, that is, t-tests or F-tests, are no longer valid. If we persist in using the usual testing procedures despite the fact of heteroscedasticity in the regression model, whatever conclusions we draw or inferences we make may be very misleading. An informal method for detecting heteroscedasticity is the graphical method. If there is no a priori or empirical information about the nature of heteroscedasticity, in practice one can do the regression analysis on the assumption that there is no heteroscedasticity and then do an examination of the residual squared e_i^2 to see if they exhibit any systematic pattern. Although e_i^2 's are not the same thing as ϵ^2 , they can be used as proxies especially if the sample size is sufficiently large. To carry out this informal method, one can simply plot e_i^2 against either Y_i or any of the explanatory variables.

Sparse Regularization

As known and stated in [Goodfellow, Bengio, and Courville \(2016\)](#), Regularization is one of the central concerns of the field of machine learning. By definition, Regularization is any modification we make to a learning algorithm that is intended to reduce its generalization error but not its training error. The no free lunch theorem (a reference in [Ho and Pepyne \(2002\)](#)) has made it clear that there is no best machine learning algorithm, and, in particular, no best form of regularization. Instead, we must choose a form of regularization that is well-suited to the particular task we want to solve.

In general, a data-based model must generalize well in order to be useful. Models that generalize well can also work consistently for new observations. On the contrary, models that do not generalize well produce inaccurate incorrect estimations for new observations in practice. The generalization ability of a model can be estimated, given a dataset, by using samples that are not used in the training process. If the model performs equally well on the training set and on the hold-out set it can be assumed that the model generalizes well. As stated in [Kronberger \(2010\)](#), the number of different states of the system must be limited and the training set and the hold-out set must be a representative sample of the whole population of observable states of the analyzed system.

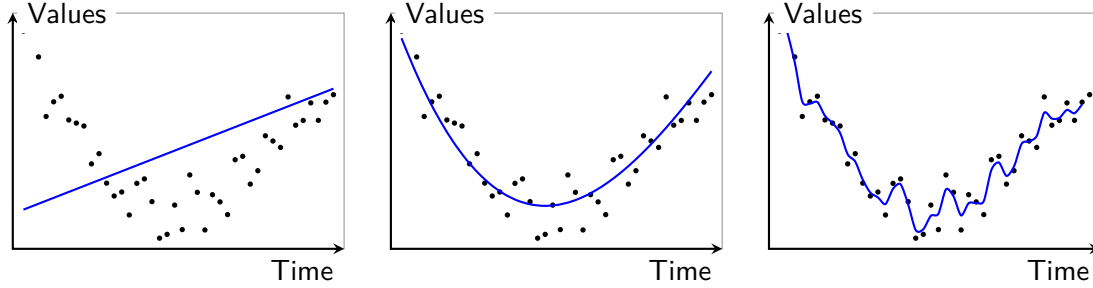


Figure 4.2: Main situations in model fitting; from left to right: underfitting, appropriate, overfitting.

Bad generalization ability is often the result of overfitting. Basically, it occurs when the model class from which a model is selected is too complex. From a complex model class, a model can be selected that fits any point in the training set very closely thus the model class has low bias. Even though the new model still fits all points of the training set very closely it has a very different overall shape thus a model class with large complexity has a large variance. In contrast underfitting occurs when the model class is too simple. A simple model class has a large bias because the model cannot fit every point closely but low variance because if any point in the training set is changed the resulting model is not very different. This is the well-known bias-variance trade-off of statistics. Minimizing the bias to get a better fitting model on the training set inevitable leads to higher variance and worse generalization. Thus the model class has to be selected in a way to find a good balance as shown in Figure 4.2. In real-world scenarios such as the one analyzed in this dissertation, bad generalization can often be the result of incomplete data. The assumption that the training set is a set representative for the whole population of possible observations often does not hold. If the training set does not contain examples for all system states that are likely to be observed, it is not possible to create a model that will generalize well to all new observations. A model trained from an incomplete training set will not generalize well if the behavior of the underlying system changes. This problem is independent of overfitting. Even though the bad generalization is not caused by overfitting in the original definition but by incomplete data the remedies to improve the generalization are similar. Reduction of complexity to increase the bias and reduce the variance of the model class should lead to better generalization behavior in both situations and this point will be crucial for some conclusions of the work presented in my thesis.

In order to avoid the above-mentioned problems and especially overfitting, the use of a suitable regularization technique can be useful. Just an introduction to this topic is given

here to the reader to advance the importance of sparse regularization for this industrial research; a complete and accurate reference can be found in [Rish and Grabarnik \(2014\)](#). As a first assumption, consider the general fitness function of a regularization method

$$\mathcal{L}(\beta) = \mathcal{F}(\beta) + \lambda \mathcal{R}(\beta)$$

and the solution of the optimization problem $\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \mathcal{L}(\beta)$.

In this work, on several occasions that will be explained, we will restrict consideration to linear models described in the form $f(X; \beta) := X\beta$.

A common choice for $\mathcal{F} := \mathcal{Q}(\beta)$ is

$$\mathcal{Q}(\beta) := \|Y - f(X; \beta)\|^2 = \|Y - X\beta\|^2. \quad (4.1)$$

With the previous choice for $\mathcal{Q}(\beta)$, different definitions for $\mathcal{R}$ lead to different well-known regression algorithms:

- $\mathcal{R} = 0$, Ordinary Least Square (OLS);
- $\mathcal{R} = \mathcal{R}_{\mathcal{R}} = \sum_{j=1}^p \beta_j^2$, Ridge Regression (RR);
- $\mathcal{R} = \mathcal{R}_{\mathcal{L}} = \sum_{j=1}^p |\beta_j|$, Least Absolute Shrinkage and Selection Operator (LASSO).

The larger λ , the smaller the “complexity” of the selected model (the variance of the estimator), at the cost of worsening the performances on the training set $\{X, Y\}$ and introducing some bias; in practical applications, λ is often used as a “tuning knob” controlling the bias/variance trade-off. The $L1$ penalization term leads LASSO coefficients to be sparse and this is a key feature for the problems at hand in this dissertation because it permits variable selection which is essential to discover which predictors are really useful among those available for the development of SSs of interest. This extremely convenient property of the LASSO allows for the creation of low-order models even when the input space has high dimension: intuitively, such a model is able to improve the stability of the predictions without sacrificing their precision. For both RR and LASSO, the best value of λ is usually obtained via cross-validation as will be explained in the core part of the thesis. The LASSO method has two main advantages over the standard linear regression as reported in [Rish and Grabarnik \(2014\)](#): (i) the $l1$ -norm regularization helps to avoid over-fitting the model on high-dimensional but small-sample data, and (ii) it facilitates embedded variable selection, i.e. finding a relatively small subset of relevant predictors.

A visualization of the sparse modeling practice for multiple linear regression problems

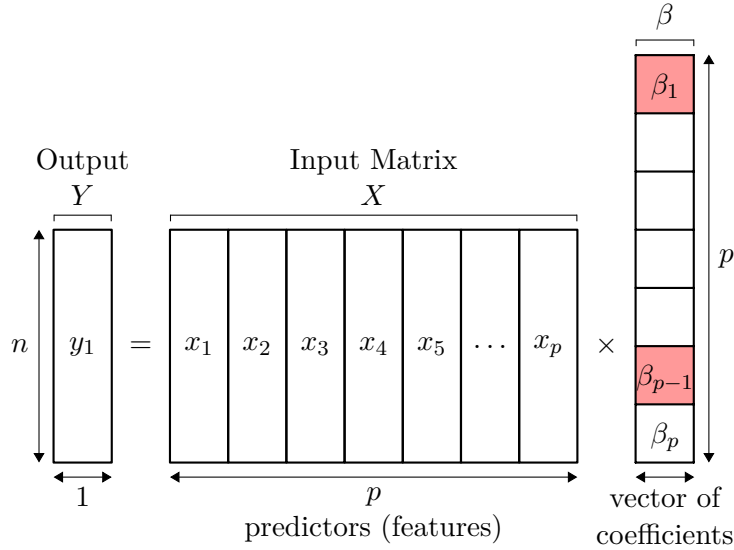


Figure 4.3: Sparse modeling visualization for Multiple Linear Regression problems in this dissertation. The highlighted β are the nonzero coefficients.

treated in this thesis is depicted in Figure 4.3 where β vector contains only few nonzero values (highlighted in the figure); these values comes from the use of sparse regularization methods like LASSO. The use of this well-known method to construct some models exploited in this thesis is also corroborated by other recent works, such as Al-Obeidat, Spencer, and Alfandi (2018) which exploits regularization in simple linear regression models to improve the performance of a model used for the prediction of temperatures within buildings.

Another possibility related to the LASSO as penalization, is the use of the *group* LASSO (as discussed in Rish and Grabarnik (2014)), which imposes an additional structural constraint besides the basic sparsity: it assumes there are groups of predictors that are naturally associated with each other, and thus must be included into (or excluded from) the model simultaneously. Without going into details here but explaining only the idea behind the approach: predictors are predisposed in groups with some kind of criterion, for example, the correlation level between them, then the l_1 -norm encourages sparsity across the groups, i.e. selection of a relatively small subset of groups from, and setting to zeros all coefficients in the remaining groups. In summary, the group LASSO works as a standard LASSO at a group level, and the sparsity parameter λ controls how many groups are selected to have nonzero coefficients in the model. The advantage of this practice instead of standard LASSO in this dissertation is the possibility to perform a selection between groups of predictors in a multiple linear regression problem and it is appreciable when data-driven approaches are exploited to improve the knowledge of a

certain process as in this case. As a matter of fact, the use of the methods discussed in this thesis have in common the purpose of advance the comprehension of the main factors involved in a precise determination of the quantity of interest (laundry weight and laundry moisture during drying), i.e. which of the available information (potentially groups of signals) on a machine is crucial to obtain SSs to control the process.

4.2 The Symbolic Regression method (SR)

A significant part of this dissertation is dedicated to the research aimed at the development of SSs for laundry moisture during drying. The use of linear models as a simple first step for this purpose, and motivated by the current limitations on firmware allowed computational effort, precludes the possibility of discovering more suitable kinds of model structures for the problem at hand. In order to overcome this limitation, a different approach was used to address the development of a SS for laundry moisture as reported in Section III. Such a method is called Symbolic Regression and exploits the power of Genetic Programming to select combinations of inputs for the estimation goal. A short introduction is provided here to the reader as a preview and to clarify the state of the art of this practice. Reporting and Paraphrasing explanations by Vanneschi and Poli (2012), Genetic programming (GP) is an evolutionary algorithm to find computer programs that solve a given problem when executed. In contrast, to genetic algorithms and evolution strategies that directly evolve solution candidates for the problem, the individuals in GP are computer programs that can be interpreted or executed to solve the original problem. The structure of GP solution candidates is not fixed, in particular, the length of programs evolved in GP is not predetermined.

This is a major difference to other evolutionary algorithms which use fixed-length solution encodings for instance real-valued vectors. Solution candidates in GP are most often encoded as symbolic expression trees that represent computer programs. The set of allowed symbols in a tree and the evaluation of trees are problem specific, and can be

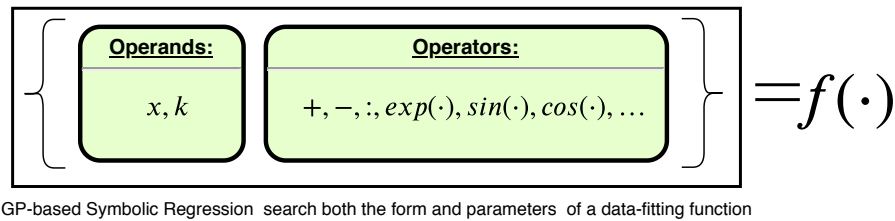


Figure 4.4: Representation of GP-based Symbolic Regression; y =laundry moisture content during drying cycle, x =independent variables (signals available), k = parameters.

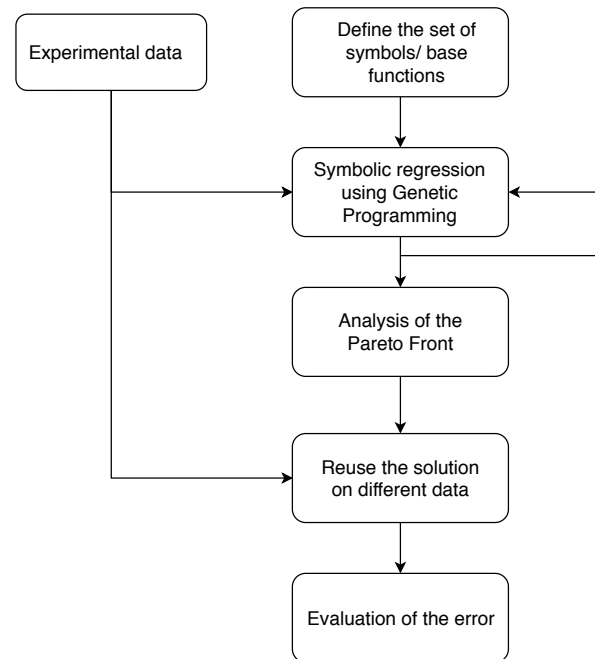


Figure 4.5: The knowledge workflow in Symbolic Regression procedure.

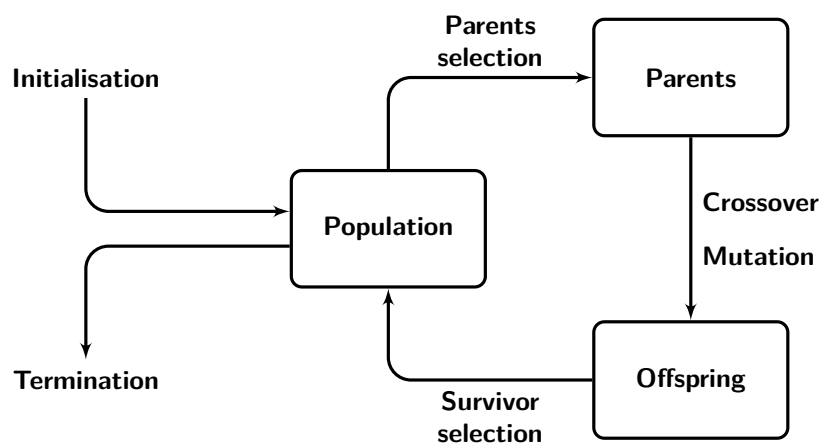


Figure 4.6: Genetic Programming workflow.

adjusted through algorithm parameters. These parameters are the *function set*, *terminal set*, and the *fitness function*. The function set contains symbols that can be used as internal nodes of the tree and the terminal set contains symbols which are allowed as terminal nodes. Generally genetic programming can be used to evolve solutions for a given problem, if it is possible to define a language describing possible solutions and a fitness function for such solutions. A representation of a general GP algorithm is depicted in Figure 4.6. The interest in this part of the thesis is to convey the main idea behind the algorithm while implementation details will be provided in 8.1. In the GP approach a population of computer programs is developed, then the idea is to move ahead on the process only the “best” models which are the ones that better adapt on the case study. The main steps of the algorithm are described below:

1. At the first step an initial population of models/programs (combining operands and operators) is randomly created (Initialization);
2. each program is executed and its fitness ascertained, then a certain number of programs from the population (parents) are selected with a probability based on fitness to participate in genetic operations;
3. new individual programs (offspring) are created by applying genetic operations with specified probabilities. The primary genetic operations are:
 - **Crossover:** a child program is created from two selected parent programs by combining chosen parts randomly.
 - **Mutation:** a new child program is generated altering a randomly chosen part of a selected parent program.
4. the best-so-far individual (program) is returned;
5. the process iterates until an acceptable solution is found or some other stopping condition is met (e.g., a maximum number of generations is reached).

One possible problem that can be solved by genetic programming and that is of interest in this dissertation is Symbolic Regression. Symbolic regression is concerned about finding a model $f(x)$ (functional expression), which is a good approximation of the values of the target variable Y given a number of input variables X . Some details about this method are introduced here to the reader.

Symbolic Regression was considered suitable for the research question of this thesis because it allows the determination of both the structure and parameters of prediction models exploited in this work as SSs. The input for symbolic regression is a dataset with n observations of the studied system and in this context, an observation is a complete laundry drying test from the laboratory as will be explained in Section 5. The result is a functional expression $f(x)$ encoded as a symbolic expression tree. For symbolic regression, the GP function set usually includes the arithmetic operators (+, −, *, /), and the terminal set includes all allowed input variables and constant values. The reader can see an example of a symbolic expression tree in Figure 4.7 where signals (operands) are combined together with operators to form the symbolic expression. A possible and frequently used fitness function is the mean of the squared errors (MSE) of the approximated values calculated by the symbolic regression model $f(x)$ and the actually observed values Y . The task of symbolic regression is a very constrained and simple task for genetic programming. Symbolic regression is thus often studied in the literature as a test-bed for new operators and other extensions of GP.

As reported in Kronberger (2011), Symbolic regression is especially useful when only a little information about the studied system is available because GP simultaneously evolves the necessary structure and parameters of the model. This is an advantage compared to other regression methods where the model structure is often fixed and only the parameters of the model are estimated. Overfitting can also occur in symbolic regression Vanneschi, Castelli, and Silva (2010) and it has been shown in the literature that this approach might lead to overfitting models that do not generalize well to new observations (for example in Kronberger (2011)). Now, as a preliminary hypothesis, based on what has been said about overfitting and simple structures, compact symbolic regression models that describe a functional relationship between the input variables and the target variable are unlikely to overfit the training data. However, it has already been shown that very compact solutions can also be less robust concerning the generalization behavior as has been stated in Domingos (1999). For this reason, one approach that has been used to reduce overfitting in symbolic regression is to use an internal validation set on which all solutions of the population are evaluated. This practice was adopted in the research described in this thesis and it will be described in detail in Section 8. As represented in Figure 4.4, the function set of genetic programming can be adapted to accommodate all kinds of different function symbols. Generally, it is not recommended to use very large function sets considering that using a restricted function has the benefit that the final solutions are easier to interpret than with comprehensive function sets. Another benefit of arithmetic function sets is that models containing only arithmetic

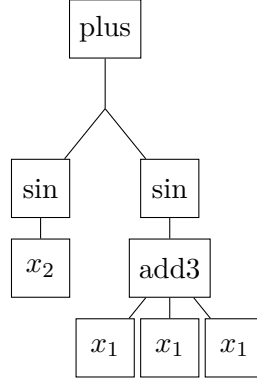


Figure 4.7: Example of a tree to represent the symbolic expression $\sin(x_2) + \sin(3x_1)$.

operators can be easily simplified symbolically. As said, one way to reduce overfitting is to use an internal validation set on which all solutions of the population are evaluated (as also suggested in other works such as [Winkler, Affenzeller, and Wagner \(2009\)](#) and [Winkler \(2008\)](#)). The problem of this approach is that without parsimony pressure the set of good solutions on the validation set is more likely to be extended at the end where the larger solutions are located. In practice the approach of using a validation set has turned out to be a working method to reduce the chance of choosing an overfit model as stated in [Winkler et al. \(2009\)](#) and that's the reason why it was considered a good practice also for the case study in my research where the problem is the development of SSs for laundry moisture prediction. As will be reported in this thesis, an indicator for overfitting is when the average or best validation fitness decreases, while the average or best training fitness is increased. This problem emerged in the use of Symbolic Regression for the goal of this thesis and it was addressed considering the correlation between input data avoiding the use of redundant information to build models and limiting the size of the input set to observations of the same kind (as explained in Section 8).

4.3 Functional Data Analysis (FDA) overview

A significant part of my research was dedicated to the study of the problem of interest, i.e., the development of SSs for drying technology applications, exploiting the entire information inherent in available signals of the process (drying cycles). The functional data analysis (FDA) tools have played an important role in this perspective. In several fields, sampling can be done on such a fine grid that it makes sense to view each sample as being observed on a continuum and coming from a smooth function. The continuum is often, but not always, time; and the functions are often, but need not be, univariate.

FDA methods were first introduced in the seminal paper by Ramsay and Dalzell (1991) and it is still a rapidly developing field. There currently are many applied and theoretical monographs available; including Ferraty (2011), Ferraty, Laksaci, and Vieu (2006), Horváth and Kokoszka (2012), Ramsay and Silverman (2007), Ramsay, Hooker, and Graves (2009). Paramount to any FDA, is that the underlying functions are smooth, i.e. that one or more of the functions' derivatives exist. The smoothness of the functions is the key property that makes functional data methods advantageous over-treating the data as discrete and using tools from multivariate statistics. The FDA tool that will be the focus of this dissertation is that of using the sampled functions in a regression model in order to understand the relationship between the functions and some other variable(s) of interest. Methods are available for when either one or both of the response and predictor in the model are functions. We will concentrate on the case of predicting a scalar response when the predictors are functions.

The basic idea behind FDA is to express discrete observations arising from time series in the form of a function (to create functional data) that represents the entire measured function as a single observation (an example in Figure 4.8), and then to draw modeling and/or prediction information from a collection of functional data by applying statistical concepts from multivariate data analysis. In doing so, it has the advantage of generating models that can be described by continuous smooth dynamics, which then allow for accurate estimates of parameters for use in the analysis phase, effective data noise reduction through curve smoothing, and applicability to data with irregular time sampling schedules. Ramsay and Dalzell (1991) present several practical reasons for considering functional data:

1. smoothing and interpolation procedures can yield functional representations of a finite set of observations;
2. it is more natural to think through modeling problems in a functional form;
3. the objectives of analysis can be functional in nature, as would be the case if finite data were used to estimate an entire function, its derivatives, or the values of other functionals.

Müller (2011) has recently described important characteristics of FDA. The FDA approach is highly flexible in the sense that the timing intervals for data observations do not have to be equally spaced for all cases and can vary across cases. Importantly, FDA methods are not necessarily based on the assumption that the values observed at different times for a single subject are independent. Although functional data themselves are not new, a new conceptualization of them has become necessary because of the increasing sophistication

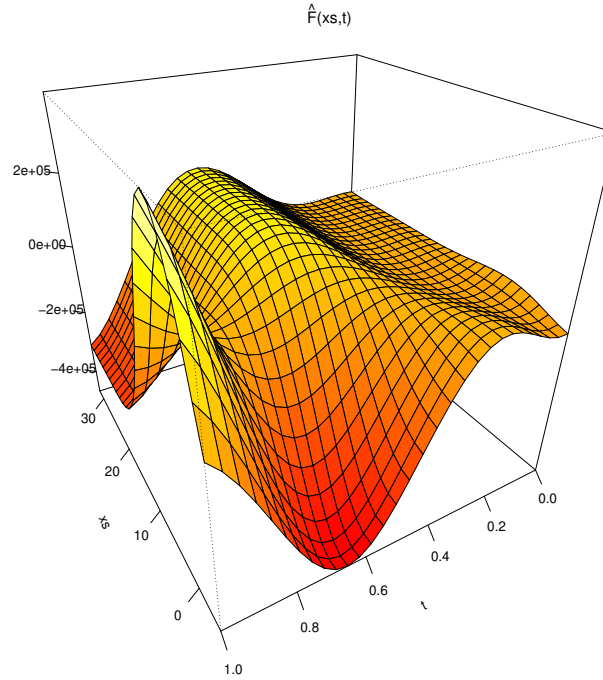


Figure 4.8: An example of the estimated surface computed using a functional regression framework. The surface describes the contribution of a certain signal x_s during time t (normalized) to prediction of the desired output (laundry nominal weight in this case).

of available data collections as stated in Ramsay and Dalzell (1991). Data collection technology has evolved over recent decades, allowing more dense sampling of observations over time, space, and other continuum measures. Such data are usually interpreted as reflecting the influence of certain smooth functions that are assumed to underlie and to generate the observations.

A full review of all FDA development exceeds the scope of this dissertation (Ullah and Finch (2013)). However, a presentation to the topic has been given here to provide an introduction to a part of the research discussed here for industrial purposes. There are many applications of functional data in industry such as Ferraty (2014), Osborne (1986) and Kalivas (1997). The present thesis can be seen as a contribution to this field by being, to the best of author's knowledge, the first application of functional techniques to the domain of fabric care major appliances. As will be discussed in the dissertation, the use of FDA techniques, and in particular Functional Regression was applied here to a case study of interest for the development of the already mentioned Soft Sensors.

Penalized Functional Regression (PFR)

In this dissertation, Functional Regression plays a significant role as an interesting application of FDA which involves the construction of functional models that describe the relation between an outcome variable and an explanatory variable. Such models are functional linear models and they are treated here as a first natural step taking into account our need for the development of SSs that are powerful in terms of predictions and interpretable at the same time. In functional linear models, the functions could be the outcome or the predictors or both and this allows new possibilities in the interpretation of current models available as will be stated in Part IV of this dissertation. Indeed, the use of functional approaches was chosen in this work to analyze a specific problem (or to construct a specific SS) which is the laundry weight estimation goal. Functional regression is also applicable to the other problems at hand, which is the laundry moisture estimation but none of the functional models introduced here have been used for this second purpose. This is mainly for two reasons:

- (i) the use of few drying cycles (observations) for the laundry moisture estimation problem is a requirement; in fact, few tests (less than 10) are often available for the calibration of the current implemented procedure based on the EoC signal. On the other hand, the use of functional methods requires the presence of a significant amount of data (for sure more than 10 observations). For these reasons, the use of functional methods was considered suitable for the other problem at hand, for which a richer dataset was made available.
- (ii) from a research perspective, the development of SSs for the laundry load estimation problem with new functional methodologies could be interesting for fabric care applications in general and not only for projects related to drying technologies. For example, a significant effort is spent nowadays to obtain more accurate predictions of laundry load during the first minutes of washing cycles in washing machines in order to perform water and energy savings. Therefore, the flexibility of FDA methods can be exploited here to solve the laundry estimation problems with more benefits than the other case study which is focused on the accurate description of a drying time series.

In the usual linear model introduced in Figure 4.1, all variables and parameters are scalars, and the regressors x_{ik} are typically assumed to be known scalars. In a functional linear model, some of these quantities are curves, and analogs of the coefficients k must be then appropriately defined. As summarized in McLean, Hooker, Staicu, Scheipl, and Ruppert (2014), functional regression models are subdivided into three broad categories,

depending on whether the responses or the regressors or both, are curves. We assume for simplicity that the responses and the regressors have mean zero. The three possible scenarios are summarized below:

1. Scalar-on-Function regression:

$$Y_i = \int X_i(s)\beta_i(s)ds + \epsilon_i \quad (4.2)$$

The regressors are curves, but the responses are scalars. The parameter is the regression function β .

2. Function-on-Scalar regression:

$$Y_i(t) = \sum_{k=1}^p x_{ik}\beta_k(t) + \epsilon_i(t) \quad (4.3)$$

The responses are curves, but the regressors are scalars. The parameters are the coefficient functions β_k .

3. Function-on-Function regression:

$$Y_i(t) = \int X_i(s)\beta_t(s)ds + \epsilon_i(t) \quad (4.4)$$

The responses Y_i are curves, and so are the regressors X_i . The parameter is the kernel (surface) β .

The main point is that the functional parameters are infinite dimensional objects which must be estimated from a finite sample. Without any restrictions on the functional parameters, a perfect fit is possible (all residuals are zero), and the resulting estimates are erratic, noisy functions which do not provide useful insights. Estimation therefore often involves additional restrictions, either by imposing smoothness conditions or restricting the action of the corresponding operators to appropriate subspaces. Several chapters of Ramsay and Silverman (2005) are dedicated to functional linear models and contain the discussion of some issues which are not addressed in this book. In this chapter, we focus on the scalar response model which bears the closest resemblance to the standard model (4.1); the discrete index $k = 1, 2, \dots, p$ is replaced by the continuous index s . This model also illustrates many issues common to all models with parameters which are functions.

Methodologies	Task 1 laundry weight SS	Task 2 laundry moisture SS
Multiple Linear Regression	✓	✓
Symbolic Regression		✓
Penalized Functional Regression	✓	
poly3 model*		✓

Table 4.1: Summary of the proposed methodologies for the design of the required Soft Sensors; methods divided by tasks, i.e. estimation goal. The (*) technique will be presented and discussed in chapter 9.

As already said, the introduction to regression techniques exploited for the research discussed here, is not exhaustive; much more details will be provided to the reader in the next sections.

The methodologies describes so far have been summarized in Table 4.1 by tasks in order to make clear the purpose for which they were used for the goals of the research. In the last row of the table there is a method that has not been described in this section being not actually a formal methodology but rather a technique that was tried combining inspiration from simple physical models and sparse multiple linear regression as will be specified in chapter 9.

A “road map” that was used as a reference for the described methodologies is the following recent work: [Fernandez-Delgado, Sirsat, Cernadas, Alawadi, Barro, and Febrero-Bande \(2018\)](#). This survey develops an empirical quantitative comparison of a very large collection of regression techniques and it is recommended to the reader interested in comparison with other similar methods not covered in this thesis. However, the last part of the dissertation presents some other methodologies which are planned for the prosecution of the research; more space will be dedicated to the methods on that occasion.

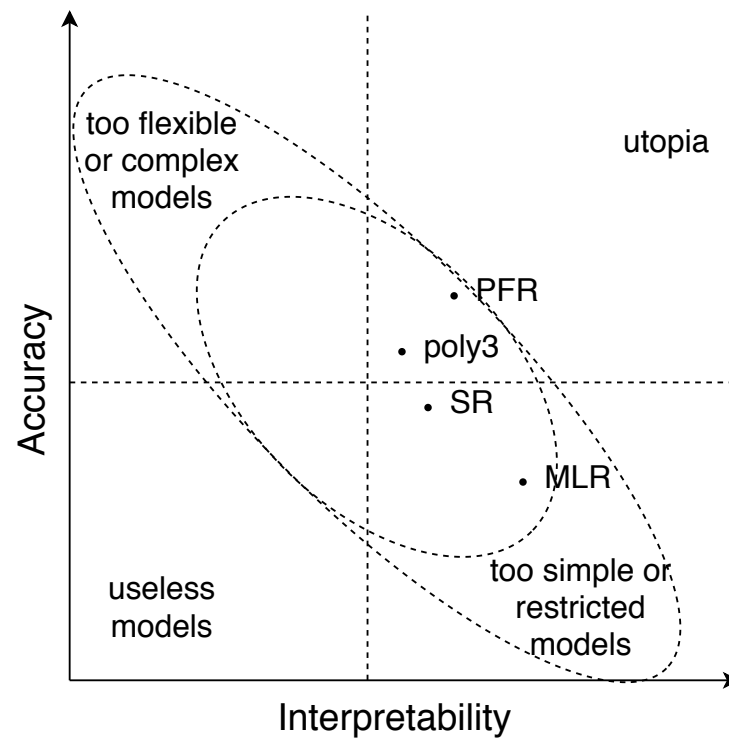


Figure 4.9: The trade-off between accuracy and interpretability of the required Soft Sensors.

4.4 A priori hypotheses on research results

The research described in this thesis is strongly focused on applications. As already stated, the SSs object of this work is thought with the aim of improving both the current drying performance on machines and current calibration procedures, in the sense that any newly introduced methodology acquires worth in case it shows advantages in terms of good prediction results and less onerous maintenance. The ways in which these requirements are measured and compared to the current set of available algorithms will be stated in parts III and IV of the dissertation.

Here the intention is to stress what the preliminary expectations are on results based on consideration of the techniques introduced so far and also on my personal experience gained in the field during these years. In this respect, having worked on several activities at the R&D Fabric care department during my PhD, I can remark, as already said in 1.5 that the main problem when some evaluations or studies are required on projects, is the scarcity of data available, in particular, data obtained under the same conditions, such as laundry composition, equipment of physical sensors, etc.

Dealing with drying technology tasks which are the object of my research, each complete drying test is always time-consuming and the quantity of different on-going projects does not allow the use of a machine or groups of machines for several days or weeks. For this simple reason, there is always the need to perform analysis with only a small amount of data available (at most few tens of complete drying tests) and this is a limit because a certain maybe significant level of variability about the trend of the acquisitions is always missing. That limit is also a requirement of this research as already said, therefore, the design of SSs designed with few data and efficient on performances is the challenging request for my Ph.D. With the intention of defining two general priorities of this project, I could state that they are

- (i) the interpretability of obtained solutions, essential for the precise comprehension of the designed SS to be implemented;
- (ii) the accuracy of the derived models from the perspective of the continuous improvement of products. From a methodological point of view, the optimal solution, considering the above-mentioned priorities, appears unfeasible taking into account our constraints on data availability.

For the sake of clarity, Figure 4.9 shows the concept in a simple way: a realistic goal for this research is the design of solutions that belongs to the inner ellipse depicted in the middle because of the requirements mentioned and the scarcity of data. All that comes out of the dashed elliptical shape could be useless or not feasible for the problem.

It turns out that a suitable trade-off between the two priorities is the real goal for my research. The introduced methodologies have been included in the representation with their acronyms and their relative position shows an a priori comparison in terms of their accuracy/interpretability for the problems treated in this thesis; the methodologies are positioned in this representation basing on the a priori knowledge of their characteristics and not yet on results, so, this kind of representation must be revised a posteriori, i.e., when results using our data will be presented.

The model which is called here “poly3” will be described in the course of the dissertation, at the moment may be interpreted as the use of a guess on the moisture model structure and (based on a fit on data with models inspired from physics) and multiple (sparse) linear regression for the estimation of the parameters of such a model. I consider all these aspects relevant to clarify which are the research hypotheses for the work at hand:

- (a) *the use of the mentioned methodologies allows model interpretability, i.e., it is supposed to be feasible to achieve a final precise comprehension of how available predictors or signals could be combined to achieve the goals. The signals that are used today by the current online implemented algorithms are supposed to be included in the set of statistically significant predictors selected by the discussed procedures;*
- (b) *with the current information available online on machines it is possible to improve the current performances without additional physical sensors;*
- (c) *the current feature engineering process involved in our already available SSs design process can be improved with the use of Functional Data Analysis techniques;*
- (d) *I am confident that also the calibration burden related to the current algorithm used for the automatic EoC in washer-dryer machines could be decreased using the data-driven SSs resulting from this research.*

These hypotheses will be discussed at the end of the dissertation (Section 14.2) in order to verify the feasibility of the research project and, in case, to provide explanations and notes about unexpected results. The next sections are dedicated to guiding the reader through the research path described in this dissertation to achieve the final research answers and realizing the related strengths and shortcomings.

Part III

Soft Sensors for laundry moisture estimation

5

Dataset description

5.1 Heat-Pump Washer-Dryer dataset

In this section, a detailed description of the dataset used for the laundry moisture estimation problem is provided. This dataset consists of a complete drying cycle (i.e., all phases described in 1.4) run on a WD-HP machine. Data are summarized in Table 5.1: drying tests are divided by the program which has been renamed here because of intellectual property rights. The target final moisture in percentage is specified for each program type together with the tolerance which is the moisture range around the target allowed for the Eoc, i.e., an automatic EoC (and therefore an estimation of laundry moisture level at the end of the cycle) is considered acceptable if it deviated within that range from the target.

Tests can be *cold* or *hot* referring to the machine temperature during the process of interest. Indeed, a WD machine could be used by the user only for drying clothes (without a previous washing cycle), and several drying cycles can be done after the first one. In all these particular cases the cycle is called *hot* because the machine temperature is “high” since other previous cycles have been executed on the machine. In the case of previous washing cycles, the entire test is called *wash and dry* (W&D), meaning that before the drying phase of the cycle there was the washing phase too. Therefore, *hot* and *cold* are terms referred to drying tests, while W&D are drying cycles preceded by washing cycles.

In Table 5.1 Type 1 tests are highlighted because Type 1 is the target load of this work: the request is the design of a data-driven SS trained with drying cycles of the program type which is most used by the user; other data shown in the table are mentioned because they are planned to be used to extend the applicability of the designed SS in the future. As will be explained below, a suitable SS for Type 1 tests could be used also with other programs changing only the final moisture target (FMT). The signals provided in the available dataset include:

- data streams of sensors;
- signals that describe the work of several components, like compressor, motor, temperature probe, etc;
- additional signals constructed from the previous ones using filters or other data manipulations techniques;
- some external information, like laboratory conditions (e.g. humidity and temperature), energy consumption, etc.

Another visualization of the data is given in Figure 5.1 where cycles are represented by their nominal weight (Figure 5.1a) and by the moisture reference/time series (Figure 5.1b) obtained from elaborations of the signal provided by a balance positioned under the machine during tests in laboratory; the steps through which the reference is derived are described below in this section. It is easy to see that moisture references are very different comparing small and large loads. A number of 56 drying tests are available for the development and 30 signals acquired for each observation.

The algorithm currently used on machines is not based on a measure or estimation of the most important output reference of a drying process, i.e. the laundry moisture content during a drying cycle. New proposals are focused on the estimation of this particular signal which is difficult or costly to measure directly; for this reason, Soft Sensors (SSs) have been developed as new solutions for the problem at hand. As regards the performance of a drying cycle algorithm, the main target is to improve the error in terms of final laundry moisture content depending on the particular moisture target.

The moisture signal used as a reference to determine the performance of the procedures discussed in this work is not a direct measure from a dedicated physical sensor (there is no ad-hoc sensor for washer-dryer machines of interest), so it is a result of some processing of the laundry weight signal which is a measure available from laboratory¹. In particular,

¹Details about the elaborations from laundry weight evolution signal will be omitted here because of intellectual property rights.

the available moisture reference is computed as a percentage according to the following relation:

$$moisture[\%] = \frac{weight - w_c}{w_c} \quad (5.1)$$

where *weight* is the laundry weight signal obtained from a balance positioned under the machine during a drying cycle, after filtering and rescaling procedures while w_c is the conditioned weight which is the weight of the dry laundry. We point out that w_c represents the laundry with the optimal moisture level for the specific program, e.g., 0% moisture level for Cotton; however also in that case there is a certain quantity of moisture within the laundry fibers and, consequently, it is possible to dry the laundry under the 0% of moisture, that's the reason why moisture references presented here can show moisture negative level percentage.

Dry Program	n° tests	FMT [%]	tolerance [%]	n° cold	n° hot	n° W&D
Type 1	56	0	±3	23	26	7
Type 2	20	12	±4	10	8	2
Type 3	15	2	±3	9	5	1

Table 5.1: Dataset available for laundry moisture Soft Sensor. FMC stands for Final Moisture Target; it changes according to the dry program type.

Each available drying test from data represents an acquisition of about 30 signals. As a first step, a correlation analysis was performed on all the tests inclusive of their acquisitions in order to discover a correlated couple of signals and possibly use a reduced set of them exploiting the same information from data. In Figure 5.2, the correlation matrix $\Gamma \in \mathbb{I}^{p \times p}$, with $\mathbb{I} \in [0, 1]$ and $[\Gamma]_{i,j} = |\text{corr}(x_i, x_j)|$, of the available signals is depicted. In general, the correlation coefficient of two random variables is a measure of their linear dependence; the MATLAB function *corrcoef* was used here to compute such values. The Pearson correlation coefficient was used for the computation (Fisher (1992)) and a level of 0.9 in the normalized scaled values was chosen as a threshold: couples with a computed value higher than 0.9 were considered correlated with implications for estimation results. In fact, in case where signals are used as predictors and correlated pairs of signals turn out to be statistically significant for estimations, only one member of the pairs will be taken into account according to this correlation analysis. The value of 0.9 together with this type of computation were chosen because of their popular use; this choice appeared suitable for a preliminary study on available data. Real names of the signals have been omitted because of intellectual property rights. Every colored square in Figure 5.2 represents a degree of correlation between a pair of signals. Basically the

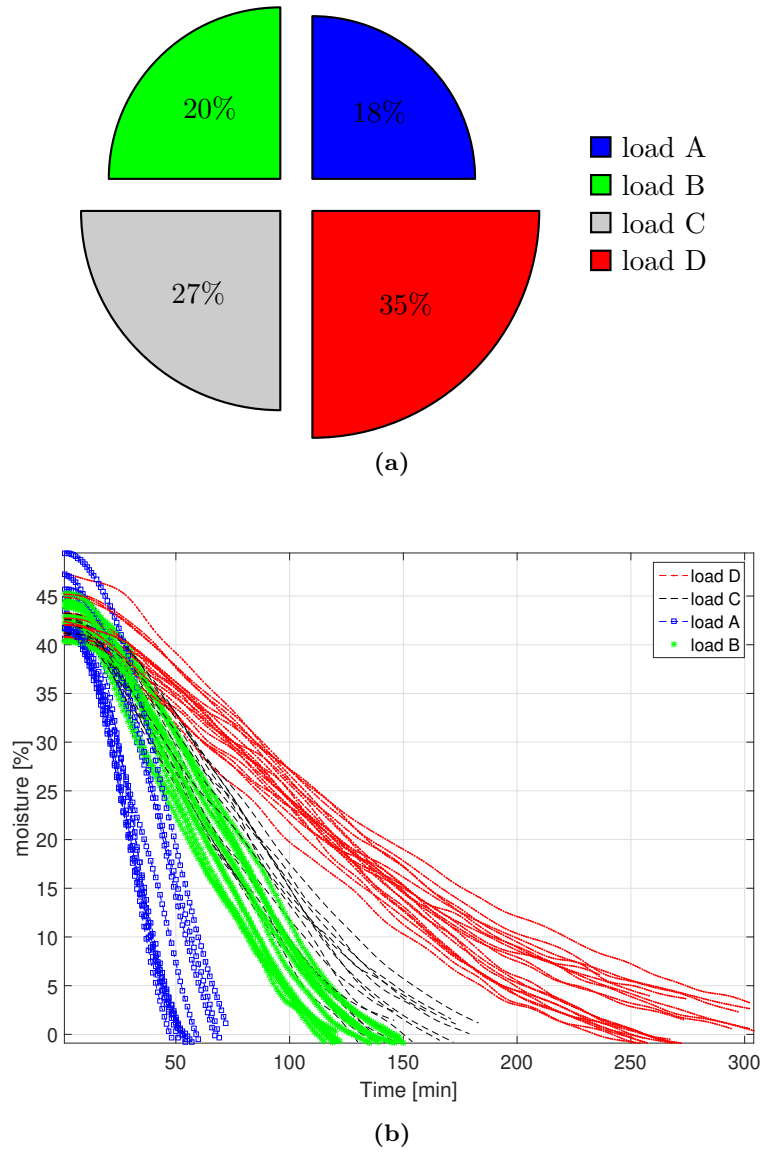


Figure 5.1: Data visualization by (dry) laundry load weight: **(a)** Laundry load weight distribution in the data, 4 classes in [kg]; **(b)** laundry moisture reference.

diagonal of $\mathbf{\Gamma}$ represents the correlations between the features and themselves, and is always equal to one. For the rest of the matrix some areas tend to have higher correlation (it happens because elaborations from the same original signals are in succession; for example the most correlated couples in Figure 5.2 refer to the same rough temperatures).

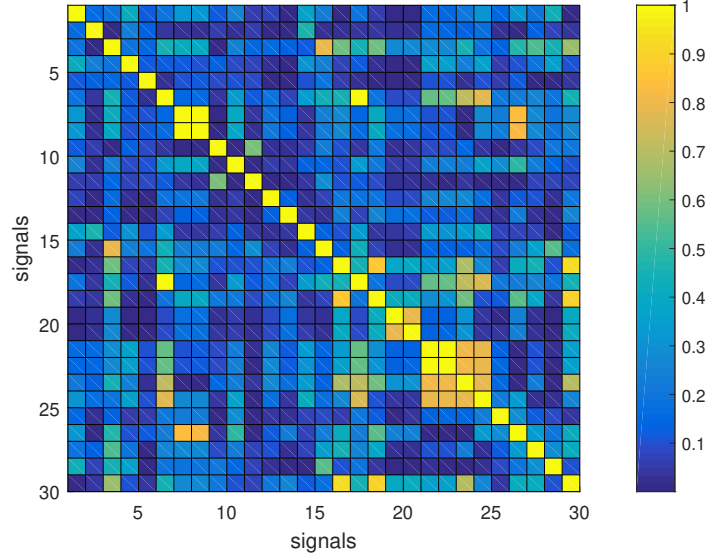


Figure 5.2: Visualization of the correlation matrix (absolute values of correlation are reported): each row and column stands for an available signal.

As already mentioned, one focus for the techniques proposed here is the automatic EoC, so, from now on all results will be reported in terms of RMSE considering the last part of the cycle (under 20% of moisture). Therefore, RMSE will always be computed considering estimated moisture ($\hat{y}_i(t)$) and reference ($y(t)$) and considering the time interval for which the reference moisture is less than 20%.

5.2 Moisture reference determination

Given that, the problem at hand is the estimation of laundry moisture content during drying cycles, a moisture reference is required for each test to train a suitable model offline in a supervised manner. Considering also the fact that the machine of interest, differently from tumble dryer machines, is not equipped with a physical moisture sensor, a moisture reference must be elaborated using the laboratory equipment. It turns out that, for this particular problem, a balance positioned under the machine during drying is a natural and simple mode to obtain the desired output that we want to use to train a model to recognize it. The rough signal from balance (shown in Figure 5.3) is treated offline to consider only the laundry weight evolution without considering the weight of the condensed water that is drained out by a pump every 10 minutes. Therefore, only a subset of the available samples is taking into account to obtain the laundry moisture



Figure 5.3: The Experimental apparatus. WD machine on balance.

reference shown in Figure 5.4a which illustrated an example of a drying cycle that belongs to the dataset. The selected samples are intentionally not reported in the figure. The selected samples are then used to perform some elaborations (a detailed explanation here is omitted because of intellectual property rights) and the results are represented in Figure 5.4b which shows the laundry moisture curve for the cycle at hand. This time series represents the desired output for the design of our SS and it is used as the reference in the supervised methods described so far. As can be seen from the figure, the drying cycle, in this case, has been extended beyond the optimal target (0% of the effective laundry moisture) using the current EoC algorithm on the machine. As an example of a trend of the current EoC algorithm on machines, Figure 5.5 illustrates the EoC signal for the same cycle analyzed so far for the reference reconstruction. Both plots in the figure represent the EoC signal but Figure 5.5a is with respect to the moisture level (Humidity) while Figure 5.5b respect to time. In this last plot it is represented also the same time sample mentioned before and related to the target final moisture (0%).

As the reader can notice the idea behind the use of the current EoC signal is the fact that its shape, for how it was constructed, tries to provide the same information carried out by the reconstructed laundry moisture seen before: the EoC signal rises up and varies more when the cycle approaches the target moisture range. The motivation for the

design of a SS for the same goal is motivated by this consideration: instead of figuring out the laundry moisture evolution with elaborations from measures, we can try to estimate the laundry moisture directly using available signals. At an early stage, the research was focused on the comprehension of the given dataset, performing the initial stages (Inspection and Selection) of the SS development scheme introduced before and summarized in Figure 2.2. As is typically the case with real data, these initial steps turned out to be intensive and time consuming for the following reasons:

- rough data usually contains acquisition problems, so the selection of suitable drying tests is always the first preliminary phase. Data visualization is essential to solve this point;
- it is always necessary to be sure about the experimental conditions: if the physical sensors on machines are exactly the same or if something changes because different machines were used. Moreover, a check on the software/firmware version is sometimes useful to understand possible different behaviors on tests performed with exactly the same hardware.

In addition, a significant analysis was made on the laundry moisture obtained from the balance since it the main reference in our supervised adopted methodologies. In particular, an attempt was made to improve such a reference retracing the elaboration of the rough weight signal. The method we used to perform this improvement is based on the observation of another curve which better reflects the drying phases of a drying process, such a curve is the drying rate and it is the first derivative of the moisture time series. We actually considered the opposite of the derivative only to obtain a signal which increases as the drying cycle goes on.

In general, a drying cycle is described observing the drying rate phases explained in the introduction for Figure 1.4. Consequently, if we have a drying rate reconstructed from an elaborated moisture, we should verify if it shows the typical shape of a drying rate curve, i.e., because of the fact that we consider the opposite of the drying rate, we must verify that it first increases, then remain approximately stable and then decreases. Moreover, all these phases must be easily visible and smooth.

With this type of consideration, the drying rate shape of the available moisture reference was represented and modified with the aim of obtaining a drying rate curve more appropriate for the description of a (laundry) drying cycle.

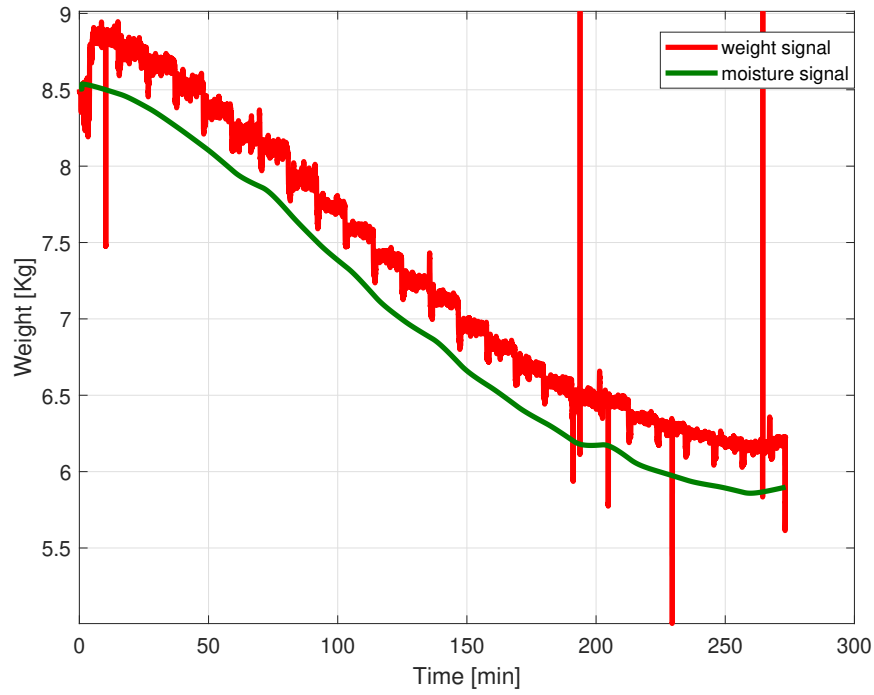
We report in Figure 5.6 an example of a result of this activity for one cycle in the dataset: the dashed line in both plots stands for the new moisture reference while the solid one is the given moisture reconstruction. The drying rate of the new reference appears closer to

the shape of a physical drying rate curve, in this example as many other cases. For this reason, the new reference curve resulting from the mentioned step was adopted for the SS design. From now on when the laundry moisture curve will be used and mentioned, we'll refer to the new reconstruction described here.

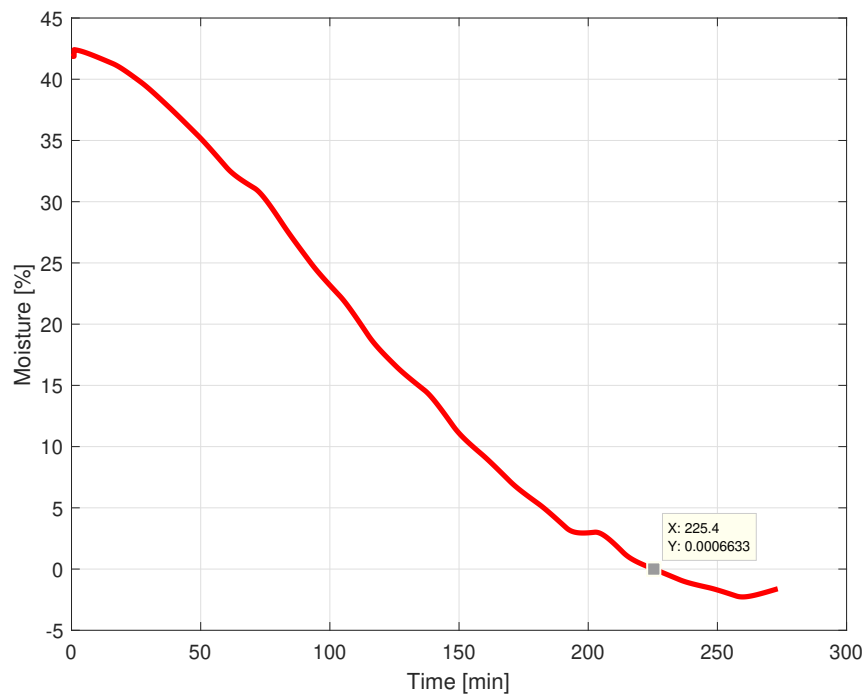
This activity aimed at the improvement of the given moisture reference involved several steps. In order to sum up, a list of the main changes is given here:

- (i) As said only some samples of the weight signals are considered and interpolated but sometimes it happens that some samples of the acquired weight signal are taken wrong by the algorithm used offline for the moisture reconstruction. For example in Figure 5.4a approximately at sample time $t = 200[min]$ the original reconstruction presents an anomalous trend motivated by the fact that, in particular conditions, the drain pump is intermittent and is activated earlier than expected. This glitch was treated modifying the reconstruction algorithm to discard these specific samples.
- (ii) another anomalous trend detected in the original reconstruction was the observation of raising samples in the moisture curve (e.g. the behavior of the moisture curve at the final part in Figure 5.4b). Indeed, a laundry moisture curve should only decrease during a drying cycle or eventually, remain stable. This problem is due to the periodic water refill of a siphon used to separate the detergent tank from the other modules of the machine. Each time the refill happens, the balance detects a slight increment. The original reconstruction algorithm treated also this situation but it was based on the acquisition of a signal that occasionally fails in detecting the condition because it occurs so rapidly. The handling of this condition was difficult because the hard problem is the distinction between these situations and different others that can lead to a steady-state behavior of the signal which is possible in a laundry drying cycle for short time intervals and in absence of disturbances.
- (iii) other minor anomalous behaviors were treated such as strange trends at the beginning of the reconstructed curve (first minutes). Also these conditions were taken into account and treated with changes in the original code. Details are not reported here because they turned out to be negligible with respect to the others discussed so far. Moreover, thinking about the goal we want to achieve, the very first minutes of the cycle is of secondary importance since our focus is on the design of a SS for the automatic EoC.

This type of activity was not planned but it appeared to be essential for a suitable setup of the required SS.

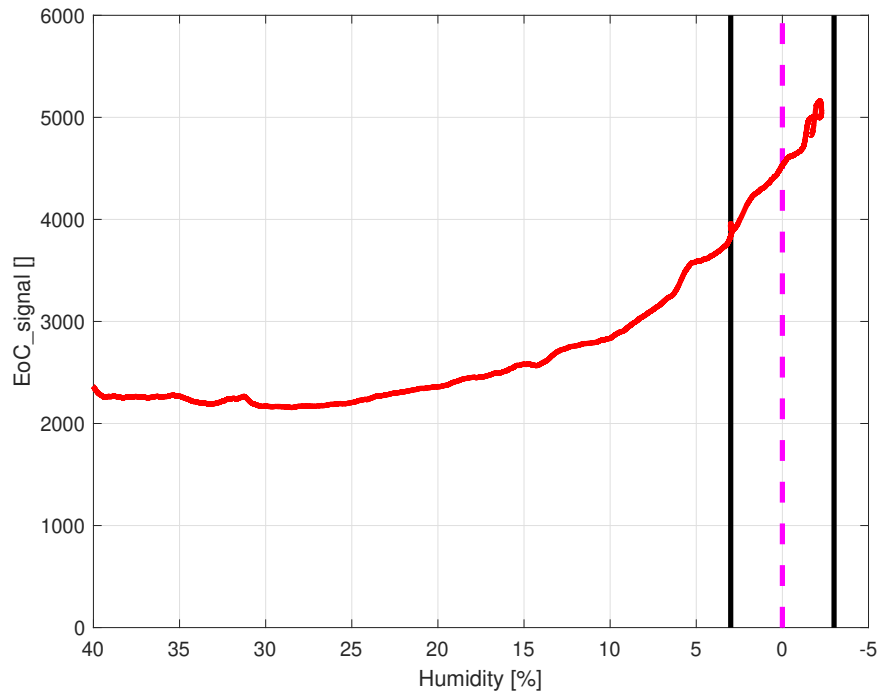


(a) Rough weight signal from balance and its elaborations. Some samples reveal acquisition losses.

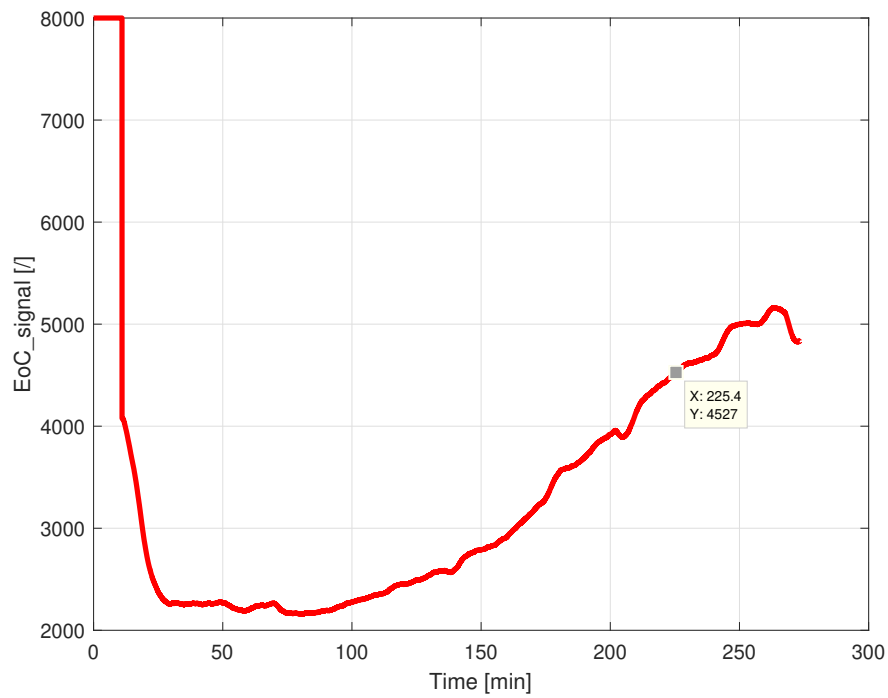


(b) Final laundry moisture reference.

Figure 5.4: An example of weight acquisition from balance and relative offline elaboration to obtain the moisture reference. In (b) also the time sample for the optimal EoC is represented (at 0% of laundry moisture)



(a) EoC signal Vs moisture (humidity); the tolerance and moisture target 0% have been represented.



(b) EoC signal Vs Time.

Figure 5.5: Examples of visualization of the EoC signal.

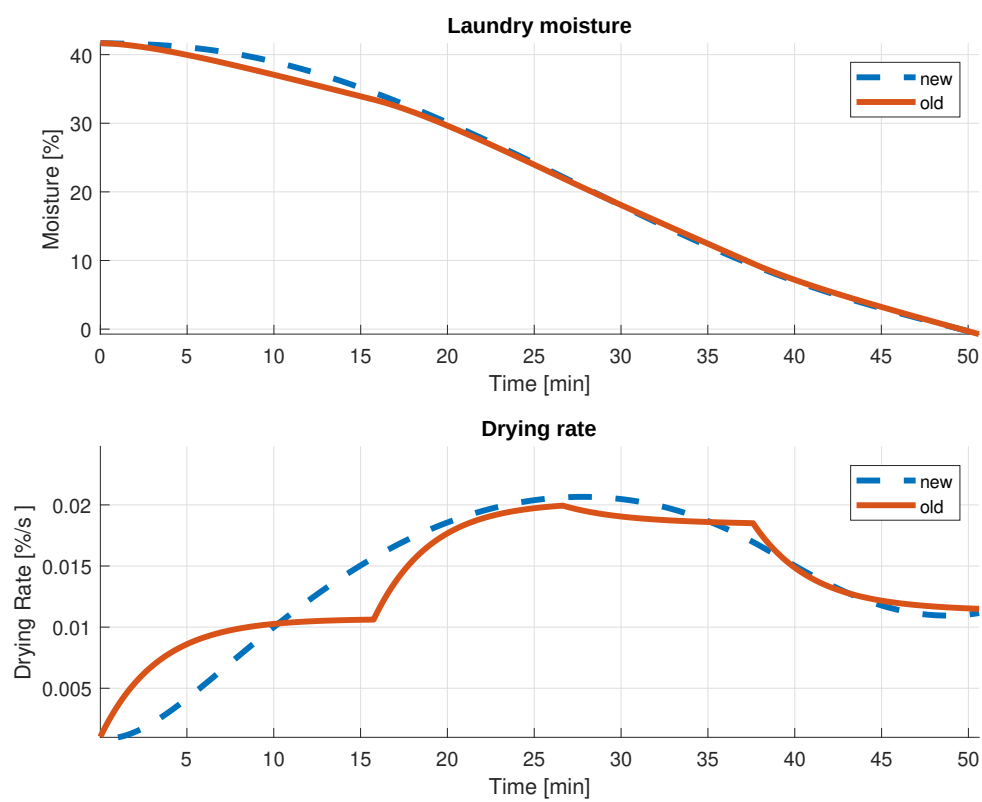


Figure 5.6: The old and new reference. Above: laundry moisture during drying; below: the associated drying rate.

6

Physical models for moisture transfer in porous media

6.1 Drying phenomena

Drying is a process of moisture transfer which is essentially driven by heat. That is, in most the practical situations of drying, the temperature field influences the moisture concentration gradient and therefore the moisture transfer process. Moreover, the moisture content influences thermo-physical parameters such as thermal conductivity, specific heat, density, and thermal diffusivity, and in some instances, it affects the material shrinkage. As a consequence, the moisture content influences the heat transfer process and therefore the temperature field within the moist material.

Although in some practical cases of drying, the heat and mass transfer processes can be modeled with sufficient accuracy in a decoupled manner, in many cases, the mathematical equations for the two processes must be solved simultaneously. Various models in the literature are available for this analogy as briefly introduced later on. The geometry of the material subjected to drying is very important and the time-dependent governing equations for moisture and heat transfer must account for it. It is of paramount importance to develop appropriate models as tools for the prediction of drying kinetics and to study the behavior of drying processes and their transport phenomena, including moisture

diffusion and evaporation used for designing new drying systems as well as selection of optimum drying conditions and for accurate prediction of simultaneous heat and moisture transfer phenomena during drying process. Since any drying process involves diffusion in the transient regime, it appears logical to treat this topic first. Here, we consider a set of simple geometries of the materials subjected to drying and write the transient diffusion equation for each treated case. Further, analytical solutions are given under reasonable assumptions for each case. However, there are two general assumptions in this section as follows:

- The heat and mass transfer processes are decoupled; that is, the temperature field does not essentially influence the moisture concentration gradient. This situation is relevant for some drying situations;
- The thermo-physical properties of the moist material and drying agent are constant.

6.2 Moisture transfer by diffusion

A drying process always involves moisture transfer which is essentially driven by heat and mass transfer phenomena. These phenomena are governed by equations similar to that of diffusion¹ (transfer along a concentration gradient), so, any drying process involves diffusion as the main physical phenomenon and equations of diffusion are always the starting point to provide a mathematical description in this case (Dincer and Zamfirescu (2016)). Moisture transfer through porous media (like laundry) can be therefore simply explained according to moisture concentration gradient variation (Figure 6.1) using the same equations exploited for heat and mass transfer (this phenomenon is known as *concentration-dependent*, Crank (1979) and Pan and Gibson (2006)).

A very brief presentation is offered here on phenomena and mechanisms of diffusion in porous media; the interested reader can find further information about the subject in consolidated literature such as: van Oudenaarden (2009).

According to the mathematical theory of diffusion, as reported in Crank (1979), the rate of transfer of diffusing substances is proportional to the concentration gradient normal to the section of the unit area. By adopting this very simple description in Figure 6.1 and assuming that the piece of laundry is representable as a slab of area A [m^2] and negligible width x [m], then, there are a total number $n(x)$ of water particles on the left of the slab and $n(x + \Delta x)$ particles on the right. The particles are randomly moving

¹In a diffusion process, matter in a system is transported from one part to another as a result of random molecular motions; there is an analogy between diffusion and transfer of heat by conduction because both of them are due to random molecular motions Crank (1979).

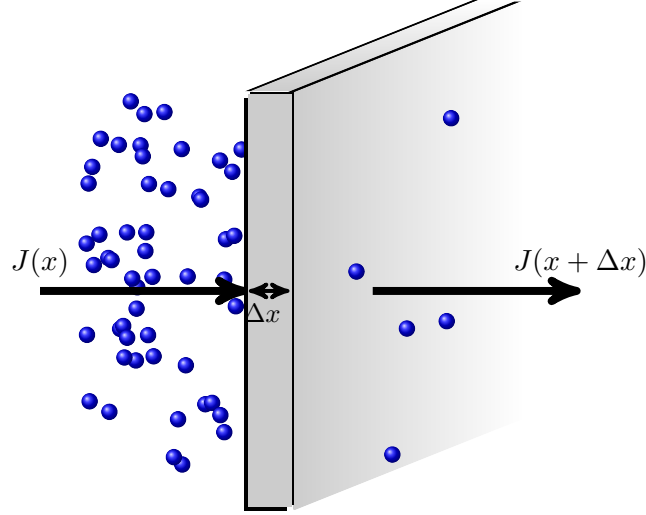


Figure 6.1: Particles randomly move through the porous medium according to a diffusion process; the arrows (vectors) represent also the flux of water particles.

and by simplifying, the probability to pass through right side or the left side is identical, so, the (positive) number of crossing to the right is $-\frac{1}{2}(n(x + \Delta x) - n(x))$. Moreover, by defining the concentration of particles as $C(x) = \frac{n(x)}{A\Delta x}$ [kg/m^3], then the flux $\mathbf{J}(\mathbf{x})$ [kg/m^2s] of water molecules through the slab during an infinitesimal amount of time τ [s] is provided by the Fick's first law:

$$\mathbf{J}(\mathbf{x}) = \frac{-\frac{(\Delta x)^2}{2\tau}(C(x + \Delta x) - C(x))}{\Delta x} = -D \frac{\partial C(x)}{\partial x}, \quad (6.1)$$

where D is the diffusion coefficient [m^2/s]. Now, assuming that the number of particles is preserved, the rate of change of the concentration during τ is as follows

$$\frac{C(t + \tau) - C(t)}{\tau} = \frac{\frac{1}{\tau}(\mathbf{J}(\mathbf{x}) - \mathbf{J}(\mathbf{x} + \Delta \mathbf{x}))A\tau}{A\Delta x} = -\frac{\mathbf{J}(\mathbf{x} + \Delta \mathbf{x}) - \mathbf{J}(\mathbf{x})}{\Delta x}$$

Considering the limits $\tau \rightarrow 0$ and $\Delta x \rightarrow 0$, it turns out that

$$\frac{\partial C(x)}{\partial t} = -\frac{\partial \mathbf{J}(\mathbf{x})}{\partial x} \quad (6.2)$$

and finally, combining equations (6.1) and (6.2), it follows that

$$\frac{\partial C(x, t)}{\partial t} = D \frac{\partial^2 C(x, t)}{\partial x^2}, \quad (6.3)$$

which is known as Fick's general second law of diffusion (concentration-dependent form). In this case, the variation of concentration refers to the variation in concentration of water and ϕ ($[kg/m^3]$) can be used instead of C . Therefore, the simplest equation that can be used to describe the process of interest can be summarized in:

$$\frac{\partial \phi(x, t)}{\partial t} = D \frac{\partial^2 \phi(x, t)}{\partial x^2} \quad (6.4)$$

where t is the time in $[s]$ and x is the spatial coordinate ($[m]$) (gradient direction through the porous media). The coefficient D here is described as a constant but it could be considered as time-dependent in a more detailed description. Equation (6.4) tells that rate of change of the concentration of water particles at a particular time and position, is proportional to the curvature of the concentration. Several methods of solution for the equation (6.4) can be found in Crank (1979) where the argument is treated in detail also for more complicated cases (coefficient D time-dependent, diffusing substance not uniformly distributed and the surfaces, etc.). The proposed solutions are too complex to be implemented in the firmware of the machines of interest, and they are subject to some initial assumptions, e.g. laundry uniformity and so on. For these reasons, the solutions proposed in Crank (1979) are considered here as the theoretical benchmark for the problem at hand and they are used as references to make comparisons in terms of performance and complexity in contrast to other simpler solutions like the one proposed by Dincer & Dost (Dincer and Zamfirescu (2016)) and reported in the following.

The general model for transient moisture transfer processes was derived as a simplified solution of the previous equation by Dincer & Dost (Dincer and Zamfirescu (2016)):

$$\phi = LF e^{-St}, \quad (6.5)$$

where LF is the so-called Lag Factor (adimensional) and S is the moisture transfer coefficient $[s^{-1}]$; the drying coefficient shows the drying capability while the Lag Factor describes the internal resistance of an object during drying. This simplified solution is widely used in literature to describe drying processes. The model provided by (6.5) is the simplest description of the phenomenon of moisture transfer through porous media; its parameters have physical meaning but this model could be inaccurate depending on the application. Moreover, a more complex physical model should be used to describe the moisture transfer in laundry during drying considering the geometry of the slab and avoiding simplifications used above. Unfortunately, this is a very hard problem in household fabric-care applications because of the several possible combinations of different laundry typology and initial conditions.

For these reasons, the simple model explained before is used here as a reference and starting point for further analysis based on real data available from machines. It turns out that one goal becomes to compare similar models in a data-driven framework even though the physical meaning is not maintained in models with more parameters. The main reference for the study on physical models for diffusion processes is reported in [Luikov \(2012\)](#), it contains details about the solution of the aforementioned equation (6.4) as well.

Here it is pointed out that, for the rest of the paper the moisture reference is computed as a percentage according to the following relation: $y[\%] = (weight - w_c)w_c^{-1}100[\%]$ where y is used for moisture, $weight$ is the laundry weight (in $[Kg]$) signal obtained from a balance positioned under the machine during drying cycle, after filtering and rescaling² procedures while w_c is the conditioned weight, i.e. the weight of the dry laundry. Besides, the desired target for improvements on automatic EoC is to keep the moisture error (RMSE explained below) under the threshold of 3% for available data (Cotton laundry). The diffusion process has been treated here as the main phenomenon involved in moisture transfer through porous media like laundry. As a matter of fact, in several cases also the advection moisture transfer phenomenon should be considered in the context of porous media. Moisture advection is the transport of water vapor by air and is included as an additional term in the equation of the process³ (6.4). [Berger et al. \(2017\)](#) show the inclusion of an advective term in the model and results in the context of moisture transfer in porous building elements while [Gasparin et al. \(2018\)](#) discuss numerically efficient methods for moisture diffusion, with a reduced computational effort. [Azeem, Boughattas, Wiener, and Havelka \(2017\)](#) focused recently on a review on the liquid moisture transport behavior of fabric stating that transportation of liquid water in the fabric cannot be defined at only one condition, but a range of conditions should be measured regarding the fabric ability to transport liquid moisture.

Considering these works and others in the literature, it is again made clear that it is complex to describe the mechanism of moisture transfer in porous media in detail because accurate physical models involve a lot of parameters and further assumptions. For these reasons, the simple physical description provided here has been used as a starting point to study the process of interest and to interpret the results of data-driven models.

Significant references for the specific problem treated in this dissertation and focused on finding models to describe drying processes in porous media are [Haghi and Ghanadzadeh \(2005\)](#), [Sander and Kardum \(2009\)](#). [Haghi and Ghanadzadeh \(2005\)](#) provides an experi-

²Details about computed features, will be omitted here because of intellectual property rights.

³Details are not reported here, the reader can see [Berger, Gasparin, Dutykh, and Mendes \(2017\)](#) and [Gasparin, Berger, Dutykh, and Mendes \(2018\)](#) as recent references.

mental examination of the convective drying behavior of wool. In [Sander and Kardum \(2009\)](#) the applicability of three models for drying kinetics was investigated. One of the most significant references for the work presented here is [Hamdaoui, Baffoun, Chaaben, and Hamdaoui \(2013\)](#) which presents results of an experimental study on textile thermal drying after wet processing treatments; its importance in this context relies both on the application, which is very similar to the one treated here (drying technology for textile fabric), and on results, because the model suggested is the same proposed in this work after the analysis on data provided by the industrial partner. The main differences between [Hamdaoui et al. \(2013\)](#) and the work at hand are summarized below:

- the experimental apparatus: here we are dealing with domestic tumble dryers while in [Hamdaoui et al. \(2013\)](#) the study was conducted on industrial continuous dryers. The latter includes more advanced energy-saving features and in general more sophisticated physical sensors. In both cases, the applications are focused on textile fabric drying.
- the general objective: in [Hamdaoui et al. \(2013\)](#) a study on an adequate general mathematical models to follow the drying phenomenon was performed. In this work, the focus is also on the procedure adopted for the online estimation of the parameters of the best model selected from data. In both cases, the mathematical expression for drying phenomenon turns out to be the same.

Three approaches will be described in detail in this part of the thesis to solve the problem of elaborating a new model to predict the output of interest using data from WD-HP machines. In this perspective, practical aspects of the use of methods will be taken into account.

7

Regularized multiple linear models

7.1 Description of the procedure

This chapter aims to describe in detail the first approach we followed for the laundry moisture SSS design. The main exploited methodology which is multiple linear regression has already been introduced and explained in 4.1; here I focus on the experimental settings and implementation aspects.

Figure 7.1 depicts the main steps of the considered algorithm: a linear model is exploited to provide an estimate per second. For example, we require an estimation using input signals at the time sample represented by the dots in Figure 7.1. The idea behind the use of linear regression here is simple: try to predict each sample of the laundry moisture value using samples of the available predictors (signals as they are) using a linear model. So, in order to make it clear:

- *inputs*: samples of the available signals on the machine. There's no feature extraction here; signals were used as they appeared in the data after the preprocessing phase which involves the resolution of acquisition problems and others as described before;
- *output*: coefficients and predicted moisture samples obtained offline from multiple linear regression using signals as predictors;

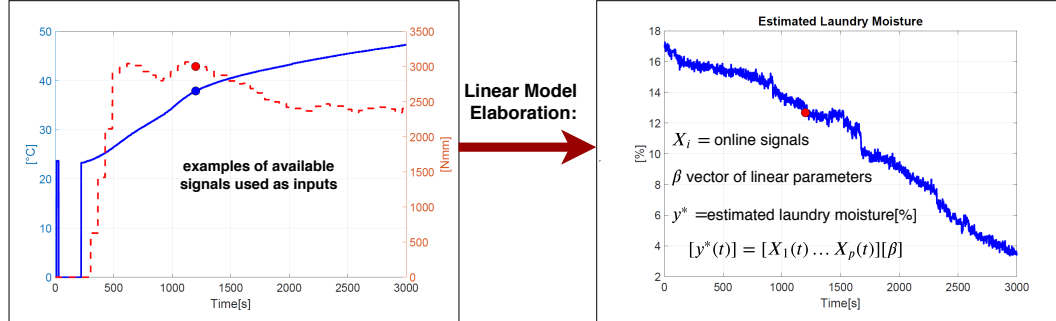
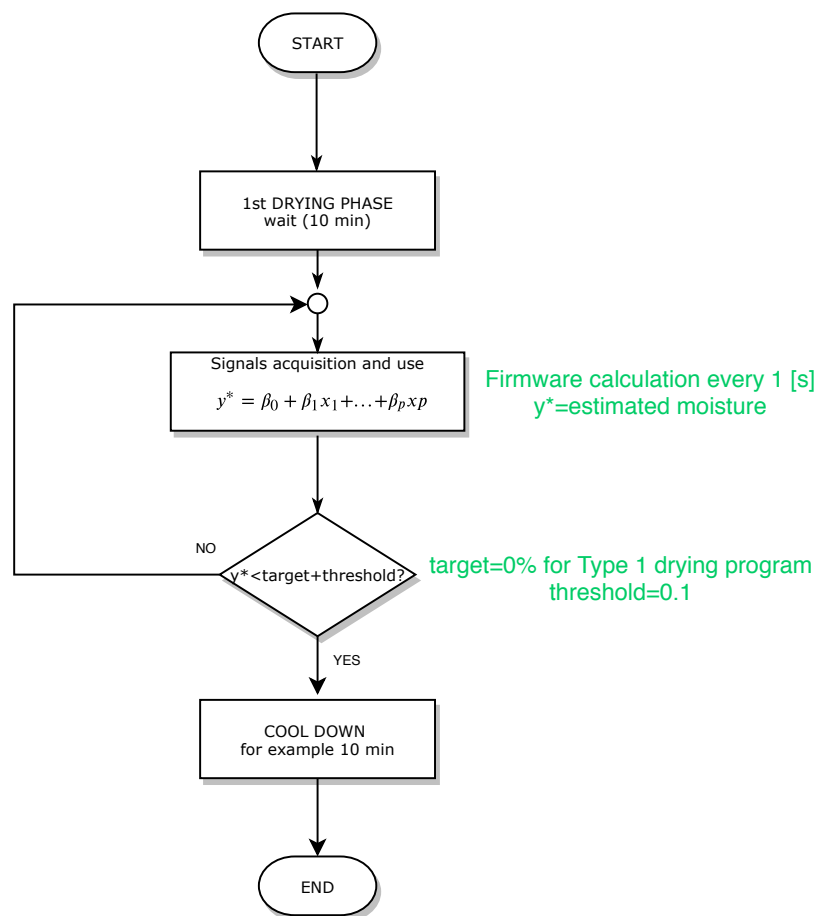


Figure 7.1: Visual explanation of the algorithm. For each time sample an estimation value is available (e.g. dot depicted in each plot).

- *reference for model training:* since we are talking about regression which is a supervised method, the reference used as the target is the moisture signal reconstructed from the weight measurement. This signal is available only offline for the SS design.

The input signals were treated offline to discover the level of correlation as seen in Figure 5.2 and to perform variable selection using embedded regularization methods such as the LASSO. Other details are reported in the next chapter. Formally, β vector is computed offline using regularization and input $(x_i(t); i = 1 \dots p)$ signals (p is the number of signals used as predictors) and an output $(y(t))$ vector of references. Figure 7.1 shows the estimation result on the right: the result is a scalar value (the dot obtained from the model and represented in the plot); a scalar output is provided for each time sample and the resulting time series is the desired laundry moisture evolution during drying.

So far, the methodology and offline implementation details have been discussed. In the interests of clarity, also the online steps are now reported here to clarify how the introduced method is thought to be implemented on firmware. This representation can work on a WD-HP which is the machine of interest for the SS at hand here but it does not change if we refer to a tumble dryer machine since the main steps are exactly the same. The best and simple way to discuss such an implementation is to start from a flow chart representation provided in 7.2. The multiple linear regression works in a single block in this scheme while the other steps have the role of ensuring the drying cycle handling: in particular a first phase after the starting point is performed to achieve the knowledge of other useful information during cycles, e.g. a laundry weight estimation works here (with some kinds of dedicated SSs that will be explained in the next part of the dissertation). Another final step is necessary to approach the termination; such a phase is called cool down and it is included to cool the machine modules after drying. During this last phase, the laundry is not dried anymore.

**Figure 7.2:** Multiple Linear Regression algorithm flow-chart.

7.2 Experimental Setting

For what has been said so far, the design of the laundry moisture SS has been made offline training a multiple linear regression model using available signals as predictors and using the improved version of the moisture reconstructed from the balance. Input and output matrices are reported here in Figure 7.4 in the interests of clarity. Y is the reference vector containing the moisture time series for each drying test of the dataset: it contains samples represented by $y_i^{(j)}, i = 1, \dots, n; j = 1, \dots, n_i$ considering $n = 56$ observations (drying cycles) in the dataset, where n_i the number of samples which is different for each observation and j is the sample index. The input matrix includes entire signals from the acquisitions and in each column, there is the sequence of the same signal for each acquisition, so, X is composed of $p = 29$ columns, where 29 is the number of available signals for each drying cycle. Using the same notation, samples in X are therefore : $x_{ik}^{(j)}, i = 1, \dots, n; j = 1, \dots, n_i; k = 1, \dots, p$.

The main implementation details are summarized here below:

- Two types of regularization were used for regression training; RR and LASSO introduced in 4.1. LASSO regularization helps to obtain sparse solutions and for this reason, it was used to perform variable selection between the available predictors. RR does not provide sparse solutions because it uses all signals to achieve the moisture prediction, but it is useful for the comparison of performance, i.e., in comparing the error obtained with the few predictors exploited by LASSO;
- For the LASSO regularizer, we made the choice of using a modified weight for the mean square error used in the LASSO optimization algorithm in order to train a model that is optimal for the EoC goal discussed in this dissertation. In simple terms: we used a LASSO that minimizes a prediction error weighted with a cost function (different for each cycle program). An example of such cost functions is represented in Figure 7.3. According to this function, the error is considered high when the moisture reference is inside the tolerance; on the contrary, it is considered gradually lower when the reference is far from the tolerance. This choice allows the construction of a model (SS) specially created for the automatic EoC;
- the result of the correlation analysis performed in the pre-processing phase (visualized in Figure 5.2) of the design was exploited to consider the input set of signals, this set has been omitted here because of intellectual property rights. Therefore, in the case of correlated couples of signals, only one was taken into account;
- Monte Carlo Cross-Validation (MCCV) (Xu and Liang (2001)) was used for training

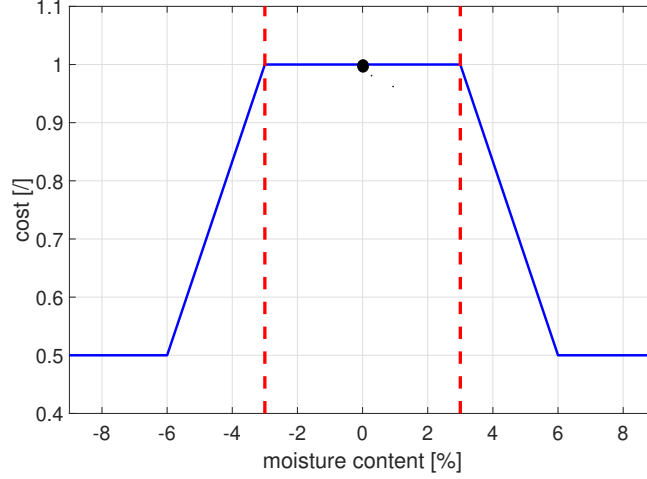


Figure 7.3: Cost function used for the error evaluation in LASSO regularization.

and testing of the designed models; in details:

- this section of the computation is made of two main cycles (or Cross-Validations): the first one used for λ hyperparameter tuning (equation (4.1)) with LASSO or RR. The second one is the Cross-Validation performed once the hyperparameter has been selected. This procedure was followed as a counter-measure for the over-fitting problem discussed in previous sections: we train models taking into account different splits of data at each iteration. Once a model is selected through the choice of λ , we test it in several scenarios picking different subsets of data.
- for each Cross-Validation loop the 70% of data were used for model training and the remaining 30% for the test. A number of 100 iterations was selected for each loop. Considering that each observation of this problem is a drying cycle that contains tens of thousands of samples, the computational burden of the entire procedure turns out to be high; for that reason, and also for the number of drying cycles in the dataset (56), no more than 100 iterations were executed considering the significance of results in terms of explained variance.

In addition, several different experimental conditions were taken into account in this part of the research. Some questions raised about the feasibility of laundry moisture SSs for specific nominal loads: i.e., train different moisture estimators offline, each one for a different nominal weight and then exploit them online making their use dependent on the knowledge of the load amount which can be given by another SS for laundry weight.

$$Y = \begin{matrix} & \left\{ \begin{array}{c} \text{moisture 1} \\ \\ \\ \text{moisture } n \end{array} \right. & \left[\begin{array}{c} y_1^{(1)} \\ y_1^{(2)} \\ \vdots \\ y_1^{(n_1)} \\ \vdots \\ y_n^{(1)} \\ y_n^{(2)} \\ \vdots \\ y_n^{(n_n)} \end{array} \right], & X = \begin{matrix} & \left\{ \begin{array}{c} \text{acquisiton 1} \\ \\ \\ \text{acquisiton } n \end{array} \right. & \left[\begin{array}{cccc} x_{11}^{(1)} & x_{12}^{(1)} & \dots & x_{1p}^{(1)} \\ x_{11}^{(2)} & x_{12}^{(2)} & \dots & x_{1p}^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ x_{11}^{(n_1)} & x_{12}^{(n_1)} & \dots & x_{1p}^{(n_1)} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1}^{(1)} & x_{n2}^{(1)} & \dots & x_{np}^{(1)} \\ x_{n1}^{(2)} & x_{n2}^{(2)} & \dots & x_{np}^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1}^{(n_n)} & x_{n2}^{(n_n)} & \dots & x_{np}^{(n_n)} \end{array} \right]
\end{matrix}$$

Figure 7.4: Visual Representation of the Multiple Linear Regression used for the laundry moisture estimation problem.

8

Symbolic models using Genetic Programming

8.1 Description of the procedure

The use of the previously discussed methodology has raised some questions relating to the use of linear models to achieve the required estimation goal. Are linear models inappropriate for the purpose? Are they too restrictive for our case study? Which kinds of different and maybe more flexible structures can outperform linear models?

The main purpose of this section of the dissertation and its relative results shown in [10.2](#) is the description of my attempt to answer to these questions keeping in mind the most important guidelines for the problem at hand and discussed at the beginning: the limited amount of available data (tens of acquisitions) and the need for model interpretability. The Symbolic Regression approach was selected as a next step in the research investigation and primarily, its use was not only finalized to improve the estimation performances, but also to reveal the nonlinear relation between input signals and output available. This proposal can be used to discover not only the coefficients of the desired model but also the structure, i.e. how some kinds of operations specified by the user could be used to process input signals to obtain the output reference. This reference is a reconstruction from the signal that comes from the balance under the machine during drying tests in the laboratory. The focus has been on the development of this method applied to the problem at hand without a priori assumptions on the model (linear or polynomial form

etc.) to discover suitable models for the purpose. Determination of an appropriate model structure is important to establish a benchmark and also to understand better the drying process itself observing which kinds of inputs are combined and how.

As introduced in 4.2, Genetic Programming (GP) is an evolutionary approach related to supervised learning. It is a domain-independent method motivated by an analogy to biological evolution. GP creates successor hypotheses by repeatedly mutating and cross-overing parts of the current best hypotheses, with the expectation to find a good solution in the evolution process. A Symbolic Regression problem consists of finding the symbolic function that matches a given set of data as closely as possible (Figure 4.4). This method is usually used to discover governing equations from noisy measurement data (Sharma and Tambe (2014)). In this work, it has been exploited to discover relations between the output (laundry moisture content) and signals already available online on the machine. A benchmarking of recent GP approaches to Symbolic Regression in the context of state-of-the-art machine learning approaches is available in Orzechowski, La Cava, and Moore (2018).

Symbolic Regression has been tested here using a GP tool called GPTIPS (Genetic Programming Toolbox for the Identification of Physical Systems, Searson (2015)) for use with MATLAB®. It employs a unique type of Symbolic Regression called multigene Symbolic Regression. Solutions found by Symbolic Regression are compact, human-readable and mathematically tractable and this makes them potentially suitable for applications like the one at hand in this work.

The use of Symbolic Regression in this research was dedicated to finding answers to the questions mentioned before and arose from the application of linear combinations to build models. Therefore, the inputs and outputs for Symbolic Regression are the same as those described for multiple linear regression as expected:

- *inputs*: available signals on WD-HP machines;
- *outputs*: structure of the model to use together with its coefficients and predicted moisture samples;
- *reference for model training*: laundry moisture time series reconstructed from balance (the improved version described in 5.2).

As a remark, it is pointed out here the fact that this kind of technique was investigated to complete the study made on the previous case with multiple linear regression while conscious of the computational effort that models constructed with these methodologies can produce. Even though a SS built with this kind of approach appears prohibitive

for implementation, it is useful to find a benchmark for the problem at hand with the available data and it helps to answer the research questions formalized before.

8.2 Experimental setting

Some choices have been made to ensure the convergence of the optimization procedure implemented in the used tool and to reduce the overfitting risk due to models which are not suitable for test data.

In particular, the configuration setting is summarized and explained:

- (i) population size = 500;
- (ii) number of generations = 1000;
- (iii) max number of genes ($Gmax$) = 5;
- (iv) max depth = 5;
- (v) number of repeated simulations = 30.

The first two points deal with the number of initial random models (*i*) and the number of allowed subsequent generations (*ii*). The choice made for these two values is essential for the convergence of the optimization part of the algorithm performed by the GP procedure; there is no specific rule to choose them, but a rule of thumb is to follow number of generations greater than the population size, typically high values in order to avoid convergence problems. Clearly, the max number of genes (*iii* – $Gmax$) is the number of trees or single models combined in the multigene approach; in this case, the value has been fixed to a low value equal to 5 in order to prevent overfitting problems using models selected by Symbolic Regression on unknown test data.

For the same reason, the max depth (*iv*) allowed for each tree has been fixed to a low value to avoid too complex models. As explained before, GP is a stochastic approach because of the randomness of the initial population of models, therefore, a fixed number of repeated simulations, i.e. a loop for the entire Symbolic Regression code has been imposed (*v*) to verify the robustness of GP to initial population. The number of iterations has been fixed to 30 in this case to avoid a huge computational effort but, obviously, it should be set to a value as high as possible. In this case, values greater than 30 for the number of iterations turned out to be critical in terms of computational effort. The stopping criterion is related to the number of generations (*ii*); a high value was used here (1000) in order to be more confident about the convergence of the algorithm. The set of operations defined by the user and exploited by the tool to try models includes: sum,

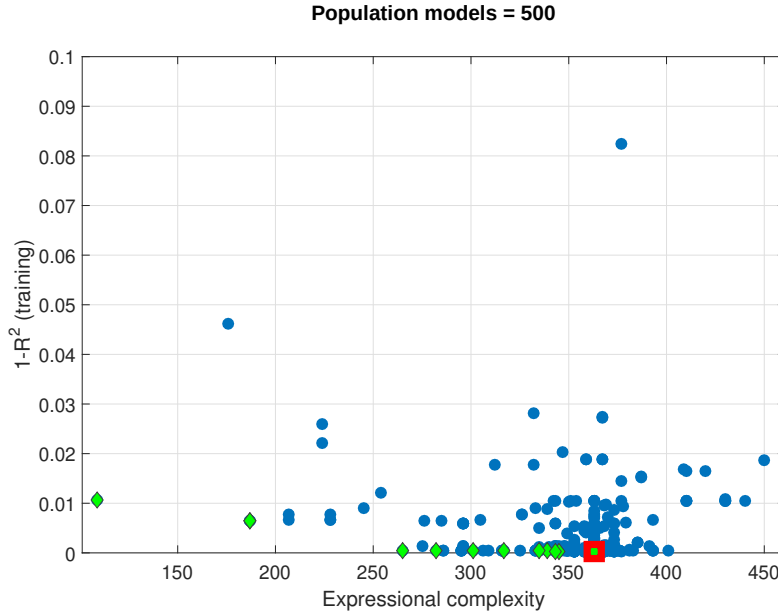


Figure 8.1: Multigene regression model population. Diamonds represent the Pareto front of models visualizing the model performance ($1 - R^2$) and complexity. Non-Pareto models are represented with dots. The best model (in terms of R^2) on the data used for training in the population is represented with a square.

subtraction, multiplication, division, exponentials, logarithms, trigonometric functions, hyperbolic functions, absolute value, square and cube functions; this set was chosen with the aim of satisfying a trade-off between flexibility and limits about the implementation burden (and taking into account the inspiration from physics - Section 6.2), but the authors emphasize that different other types of operations could be included.

The tool GPTIPS [Searson \(2015\)](#) used for this analysis helps the user to identify a model that looks promising for further investigation. In Figure 8.1 each model is plotted as a dot with $(1 - R^2)$ on the vertical axis. As said before, the coefficient of determination R^2 is the square of the correlation coefficient which is the ratio between the covariance of the two variables and the product of their standard deviations. The range is 0 to 1, so the target is to keep the difference $(1 - R^2)$ close to 0 and the complexity of the expression on the horizontal axis. The Pareto front¹ models are highlighted in green. The final result for a single iteration is determined by the best trade-off between the model complexity and the fit error (red circled dot in Figure 8.1) which is automatically selected by the algorithm.

An entire moisture time series is selected randomly at each iteration to belong to a

¹Model in the Pareto front are models that are not dominated by other models in the population in terms of model complexity and fitness (see [Lotov and Miettinen \(2008\)](#) as a reference.).

training or a test set and for each case, the Symbolic Regression selects the best model according to the steps required by Genetic Programming. Performances in terms of RMSE for drying test i are computed: $\sqrt{\frac{1}{n_i} \sum_{t=1}^{n_i} (y_i(t) - \hat{y}_i(t))^2}$ where n_i is the number of samples for each observation and $\hat{y}_i$ is the estimation in percentage of laundry moisture; error distributions are then visualized to evaluate the goodness of the provided models (see 10.2). Estimation error is always computed considering the last part of the cycle: under 20% of reconstructed moisture because it is the part of interest in the application for the EoC.

Since this methodology was chosen to investigate the possibility to use more flexible models for the purpose, the primary focus was not on implementation aspects or computational burden, but on performance and the discovery of suitable structures for the laundry moisture SS design. As will be stated later, the use of Symbolic Regression gave me confirmations about the role of certain groups of predictors involved in the estimation process but it didn't provide a winner structure able to avoid the significant overfitting related to this case study where the lack of data is a characteristic of the problem and not a circumstance as said. Flowcharts for an online implementation are not available in this case since this methodology was employed as a prosecution of the research activity conducted with linear regression as explained before.

9

Polynomial model: laundry moisture prediction as a time series

9.1 Description of the procedure

It has already been stated how the most difficult part of the design of a SS in general and also for our purpose is the data acquisition and pre-processing: inspection, outlier detection, sample selection, data visualization, etc. The model design part, once the experimental settings have been determined, is a systematic execution of the discussed methodologies. Naturally, this research is concentrated on finding the most simple solutions to achieve the goal which is the improvement of the performance of the current EoC algorithm with the limited data available. As a consequence, the first exploited methodology for the model design was multiple linear regression using input signals as they are; then Symbolic Regression was analyzed to complete the study using the same inputs. After that, the research proceeded in a different direction oriented to a method which combines the expert knowledge of the process with the requirements and what we learned from the use of the previous methodologies.

The technique discussed in this part will be called “poly3”. It was not introduced in the section of methods presented in this dissertation and the reason is that we are now dealing with a procedure that actually exploits multiple linear regression in a different

way, i.e., with different inputs and outputs with respect to the case already explored. In fact, let's make clear in which sense the case is different now:

- *inputs*: available signals from WD-HP are now not considered as they are but their information is codified and represented computing predefined values from them, the so-called “features”;. Another input is the time elapsed from the drying cycle start;
- *outputs*: in this case, the outputs are the coefficients of a moisture model which is not a combination of signals anymore but a pure time series, so a combination of time samples with coefficients estimated with the inputs;
- *reference for model training*: laundry moisture time series reconstructed from balance (the improved version) as seen for the other explained methodologies.

Concerning what has been said, the main difference, in this case, is the use of the model which describes the laundry moisture in time without being a combination of signals. One advantage of this practice is immediately clear and is simply based on the possibility to satisfy the TTE goal presented at the beginning of the dissertation. Indeed, with the methodologies described so far the estimated moisture value is available at the same sample time in which signals are acquired and combined together with multiple linear regression or symbolic regression; here the model can be determined after a suitable period of time from the cycle start (to acquire a sufficient amount of information) and then it can be used to predict the entire drying cycle. Clearly, this different approach requires a significant effort on the estimation of the coefficients of such a model and, in this regard, the feature engineering phase is crucial to perform a suitable and powerful representation of available signals.

Going into detail, the structure of the model is now fixed but it is nonlinear and determined after a preliminary offline analysis comparing some candidate forms provided by literature (models used to describe moisture transfer in food preservation and other areas different from fabric care).

Figure 9.1 shows an explanation of the proposed procedure in detail. The available dataset has been used offline to select the best model for the purpose and it was done checking fitting performance (as summarized in Table 9.1 reported below) comparing each candidate model and the available output which is the result of some elaborations from the signal of the weight evolution of laundry during drying cycle in laboratories. The best model between the candidates was chosen by considering some well-known indices used for goodness-of-fit analysis: Root Mean Square Error (RMSE) and coefficient of determination (R^2). RMSE is represented in the next Section; the coefficient of determination is the

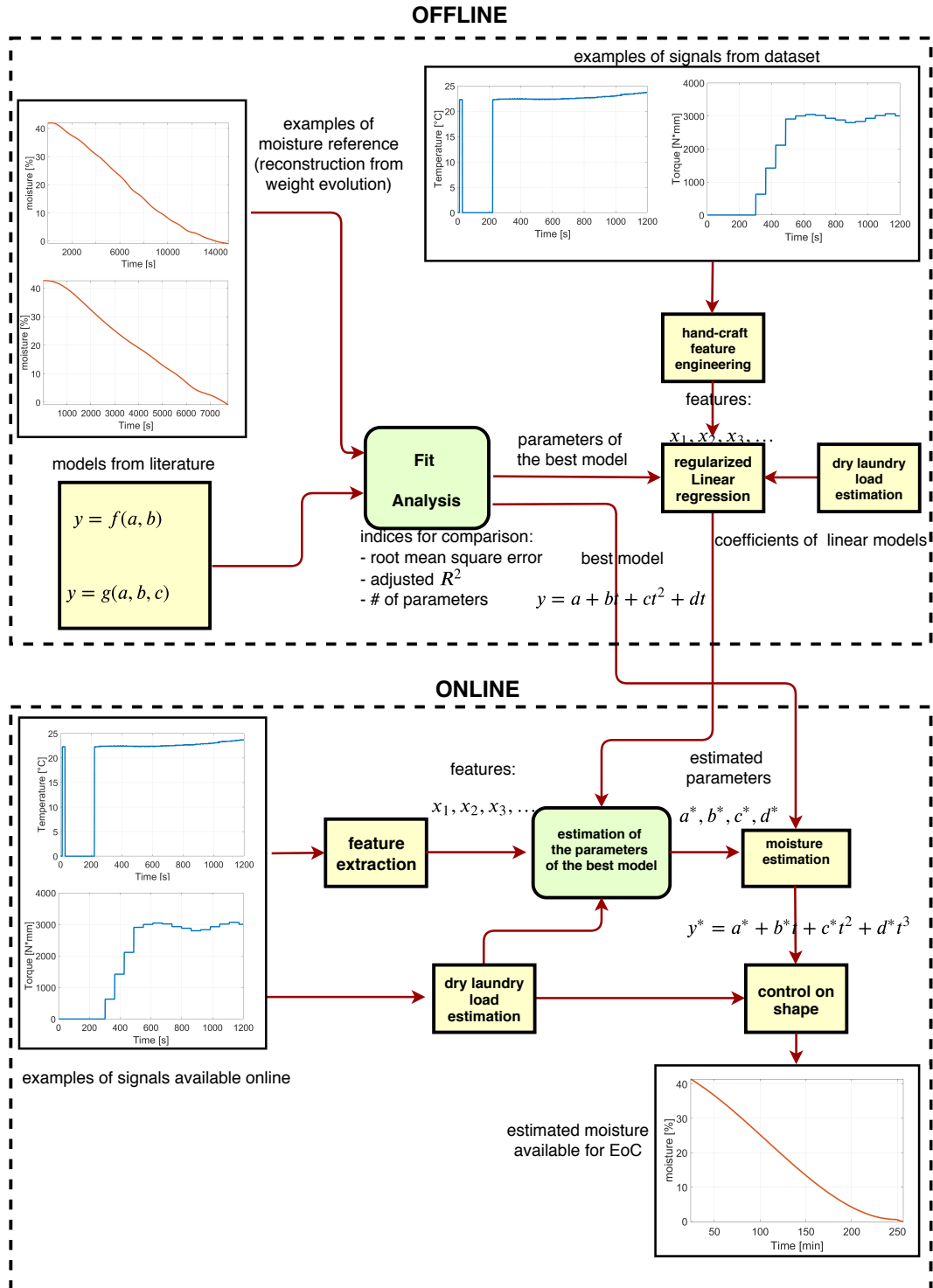


Figure 9.1: Visual explanation of the polynomial SS.

square of the correlation coefficient; a higher coefficient of determination is an indicator of better goodness of fit for the observations. Moreover, also the number of parameters for each model was taken into account in order to avoid too complex models for a firmware implementation.

The final trade-off (as shown in Table 9.1) was found in a 3^{rd} order polynomial model in time (t) and parameters a, b, c, d :

$$\hat{y}(t) = a + bt + ct^2 + dt^3 \quad (9.1)$$

This is the structure that provides the best fit performance on data (WD-HP case study) and is proposed here as the model to explain the evolution of drying cycles of interest. After fit analysis, the structure of the proposed model is fixed and the values of parameters for each observation as well (i.e. for each observation the fit analysis provides the values for a, b, c, d which are used then as the output reference for the linear regression step). Once the model and its true parameters are fixed, we need a procedure to estimate such parameters online using the available information from signals.

The method chosen here to do this is to follow a simple linear regression procedure, using some values computed from signals as inputs; these values will be called features and summarize the information provided by each available signal. These features were defined using expert knowledge and they are fixed. The hand-crafted features used¹ to estimate the parameters of the polynomial model are available after $t^* = 20[min]$ from the beginning of the cycle, i.e. after 20 minutes from the start phase an estimation of the laundry moisture content for the entire drying cycle is ready; examples of such features are quantities computed from Temperature at the output of the drum and from the motor torque during drying.

Essential information exploited here is the one related to the (dry) weight of the laundry [dry laundry load estimation block in Figure 9.1 – OFFLINE part], which is a value that is estimated online using the available information of the inertia of the laundry weight. So the information on the laundry weight quantity is used to train a linear model offline which is different depending on the laundry weight: the structure of the model is always linear but the coefficients of the trained linear model vary depending on the laundry weight level.

¹Details will be omitted here because of intellectual property rights.

Model	expression	RMSE[%]	R^2 index	#params
Lewis	$y = e^{-at}$	9.069	0.781	2
Henderson & Pabis	$y = ae^{-bt}$	1.147	0.996	2
quadratic	$y = a + bt + ct^2$	1.051	0.997	3
cubic (poly3)	$y = a + bt + ct^2 + dt^3$	0.399	0.999	4
rational	$y = \frac{a+bt}{1+ct+dt^2}$	6.564	0.897	4
gaussian	$y = ae^{-\frac{(t-b)^2}{2c^2}}$	0.417	0.999	4
sigmoid	$y = a + b\frac{1}{(1+ce^t)}$	0.496	0.999	3
two exp	$y = ae^{-bt} + ce^{-dt} + e$	0.746	0.998	5
mixed	$y = ae^{-bt} + ct^2 + dt + e$	0.453	0.999	5

Table 9.1: Drying models fitted to experimental data; t =time, y =moisture.

Table 9.1 summarises the results of fitting moisture reference models and several models used for different drying applications (see also Kucuk, Midilli, Kilic, and Dincer (2014); Haghi and Ghanadzadeh (2005)); this study is represented as the “Fit Analysis” block in Figure 9.1 – OFFLINE part.

At the end of the offline part the form of the model is available together with the coefficients of linear models used to estimate the parameters of the polynomial structure; for each parameter a linear model has been used exploiting the same set of defined features, so in this case 4 linear models are involved. The online part is executed on the machine (WD-HP) and it deals with some simple steps: the first is the feature extraction process which is performed computing the features defined in the offline part after the first phase of the cycle (20 minutes from the start). In this manner, we have time to collect useful information from signals. The second step is the merging of the computed features with the coefficients determined in the offline part and such a merging is made-up by the linear models used to estimate the 3^{rd} degree polynomial parameters. The set of coefficients used by the firmware here varies depending on the estimation provided by the laundry weight estimation procedure [dry laundry load estimation block in Figure 9.1 – ONLINE part]: once the estimation of the (dry) laundry load weight is available², it is used to select the correct set of coefficients of the linear models employed to determine the estimation of the moisture model parameters (a^*, b^*, c^*, d^*). Therefore, for each load size, a set of linear models has been trained, for example:

²Load estimation is done using linear models as well and particular features; details are not reported here.

$$\begin{aligned}
a^* &= \alpha_0 + \alpha_1 h_1 + \cdots + \alpha_7 h_7 \\
b^* &= \beta_0 + \beta_1 h_1 + \cdots + \beta_7 h_7 \\
c^* &= \gamma_0 + \gamma_1 h_1 + \cdots + \gamma_7 h_7 \\
d^* &= \delta_0 + \delta_1 h_1 + \cdots + \delta_7 h_7
\end{aligned} \tag{9.2}$$

where $h_i, i = 1, \dots, 7$ are the features selected offline and $\alpha, \beta, \gamma, \delta; i = 0, \dots, 7$ are linear regression coefficients.

It is pointed out that the computation of the drying rate which is the derivative of the moisture content in time, can be useful to detect anomalous estimation of laundry and this fact is exploited in poly3 to modify the estimation in such cases. This situation is handled as a 'control on shape' of the polynomial model (Figure 9.1). This simple control consists of a simply fixed decrease of the moisture for every predefined time interval [every minute]. When the computed drying rate goes under a certain threshold (equal to 0[%/s]) the control starts and continues until the moisture target is reached (0% in case of Cotton).

9.2 Feature engineering

This is the core part of the description of the method at hand here. As already discussed, the inputs for the designed model are features i.e., values computed from the acquired signals and that represent the information required to perform the automatic EoC determination. This part of the SS construction is included in the preprocessing phase (scheme visualized in Figure 2.2) and it is, in general, the most important step of the entire research process. In this case, it has proved to be critical and time-consuming because it required the fundamental contribution of the experts of the process and the specific machine. The feature engineering process conducted in this part of the research is *hand-crafted*, in the sense that features from signals were chosen with a dedicated activity for this purpose collaborating with the experts. This kind of activity is onerous because several skills are required to do it efficiently: for example, the experience gained over many years by the experts of the current implemented EoC algorithm on machines allowed the suggestion of which type of signals are a priori more suitable for describing the evolution of the laundry moisture evolution. Then, features, once computed, have to be studied to determine the correlation between each other in order to avoid the use of redundant input information. Another key point is the feature selection that here is

intended as the determination of the most useful features after the model design phase when features have been used as predictors for estimation (we refer again to the general scheme for SS development in Figure 2.2).

It is remarked here how the use of MLR and SR techniques helped to improve the knowledge of the process allowing us to identify also the statistically significant predictors and this was an important starting point for the feature engineering step discussed here. Because of the above-mentioned reasons, we are talking about a critical point of this research; it turns out that estimation results and the entire SS design depend on how the feature engineering process is conducted. In general, this is true whenever the research is focused on the representation of the input information to obtain a data-driven model. Paraphrasing a concept expressed in [Halevy, Norvig, and Pereira \(2009\)](#): *Good features allow a simple model to beat a complex model*; in this case, good features are the ones that captures the essential information of drying process evolution.

A part of my Ph.D. was dedicated to the improvement of the feature engineering process aimed at the construction of SS for fabric care drying technology keeping in mind on one side the requirements of interpretability and on the other side the scarcity of data available for evaluations. In fact, the use of different strategies such as deep learning could be attractive also considering its strengths with regards to the feature engineering process ([Goodfellow et al. \(2016\)](#)), but in this case, it does not appear suitable for the requirements and identified problems.

signal	name	explanation
signal 1	x1	value @ 20[<i>min</i>]
signal 2	x2	value @ 20[<i>min</i>]
signal 3	x3	value @ 20[<i>min</i>]
signal 4	x4	std in [1, 10] <i>min</i>
signal 5	x5	Max-min in [0, 20] <i>min</i>
signal 6	x6	Max value observing drum speed
signal 7	x7	std in [0, 20] <i>min</i>

Table 9.2: Features used to estimate the parameters of the polynomial model; std stands for standard deviation. Other details omitted because of intellectual property rights.

The result of the feature engineering activity described so far has been summarized here in Table 9.2 where features are described referring to the relative signals. This table is the result of a more complex list of features that were processed in a study of their correlation and finally proposed as inputs. The correlation analysis was similar to the one explained in 5.2 using the Pearson correlation coefficient and a threshold level of 0.9

similarly to the case seen in the description of the dataset. The real names of signals and other details have been omitted here because of intellectual property rights.

9.3 Experimental setting

It has been explained how the procedure used here takes into account the multiple linear regression described in previous sections. As a consequence, also some aspects of the experimental setting are shared with the characterization provided in 7.2. Linear models used to obtain the parameters a, b, c and d for the model summarized in equation (9.1) were trained using Regularized Linear Regression with LASSO which is exploited for its sparsity to select a set of statistically meaningful predictors (the features mentioned in 10.1). In this phase, 70% of available observations for each load are used as the training set, while, 30% are used for test in 100 Monte Carlo Cross-Validations.

A peculiarity of the procedure at hand here is actually the use of another type of SS used for a different goal: the laundry weight estimator. As represented in Figure 9.1, a block for dry laundry load estimation is included offline to train models and online to select the correct model to use. This block was included to improve the precision of the desired SS having assessed problems related to the variance of obtained estimations. The rationale behind the use of this practice in the algorithm is simple: we develop SS for the estimation of similar moisture profiles because they belong to the same group of nominal load; the model is trained to recognize time series that are similar in terms of duration and this helps to limit the model variance using few data. In these terms, several moisture models are prepared offline while on the machine the laundry estimation block will provide an indication of the laundry loaded in the drum and so it will select which moisture model is suitable for the current running drying cycle. The laundry weight estimator mentioned here is simply based on the inertia measures presented in 3.1, Figure 3.5. The use of a SS (like the one described in Figure 3.8) dedicated for this purpose is under evaluation also in view of the results we are achieving using a different methodology as will be explained in part IV of the thesis.

In order to make all these aspects more clear, a flow-chart of the procedure thought for the implementation has been represented in Figure 9.2, similar to the one depicted in Figure 7.2 for multiple linear regression method as regards the initial and final part of the drying cycle (cool down phase).

Another implementation detail that arises from the visualized scheme is the one summarized with the block called “control on shape”. It consists of a simple control on the moisture estimated values provided by the algorithm because, in general, a drying

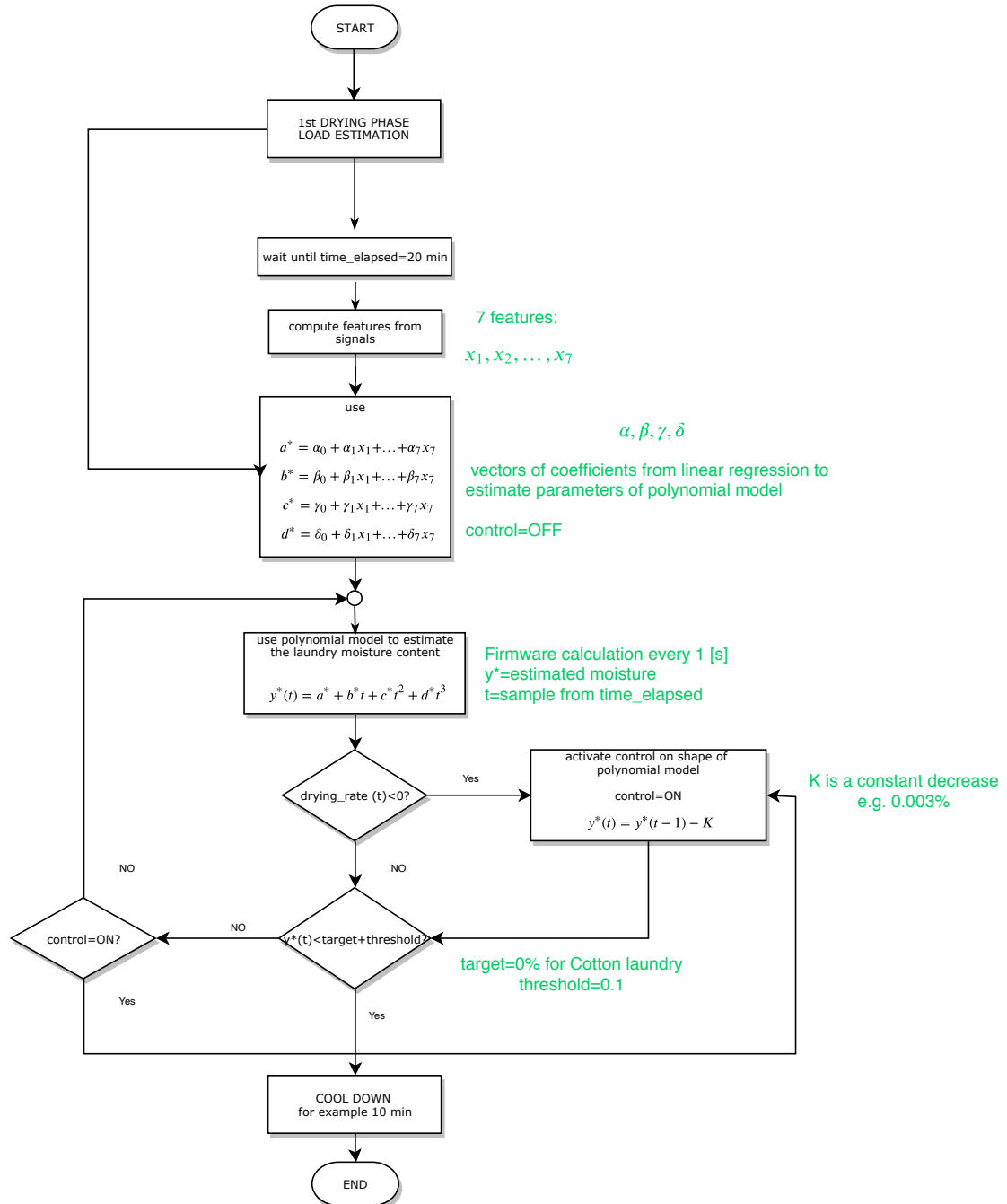


Figure 9.2: Algorithm flow-chart for the automatic EoC using polynomial model in time.

curve cannot increase during a drying cycle (as explained in 5.2). With this in mind, in case of an estimation that presents an increment of the moisture, these conditions are detected through the drying rate observation which is the derivative of the moisture in time computed with few samples (10) together with the estimated moisture once available. In case of this anomalous condition, a simple control is activated and it consists of a constant decrease of the estimated moisture. The value for a constant decrease was chosen from the observation of the results; it is suitable to avoid best-case scenarios considering the available programs and laundry weight decided by the user.

10

Experimental results

10.1 Regularized linear models

In Figure 10.1a the main results in terms of error distribution are depicted. The estimated moisture value is provided for each $t=1[s]$ by the model and the estimation error is computed using the moisture reference that comes from the laundry weight signals available from data as said before. The represented error is the Root Mean Square Error (RMSE) and in particular, from left to right, an increasing size of input signal has been used exploiting the sparsity of LASSO and imposing the input dimension as a degree of freedom in the code, e.g. LASSO1 stands for linear regression which uses lasso regularization with n_s signals as inputs at most (n_s is a predefined value which is omitted here); so, for each iteration, 30 signals are available as input but a number of signals (predictors) less than (or equal) to n_s are selected by the algorithm because n_s is specified as the maximum number of nonzero coefficients in the models tried.

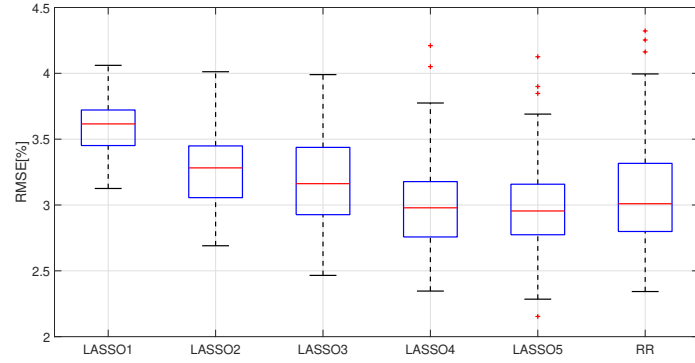
The error distributions of all the simulations performed, in this section as well as throughout the thesis, are reported in the plots using boxplot representation. In particular, boxplots are used as visualizations of distributions of the cross-validation errors taking into account the test set of the available data over Monte Carlo iterations (100 unless otherwise specified) for a given set of hyper-parameters.

The median RMSE of the distribution represented by each boxplot in Figure 10.1a is

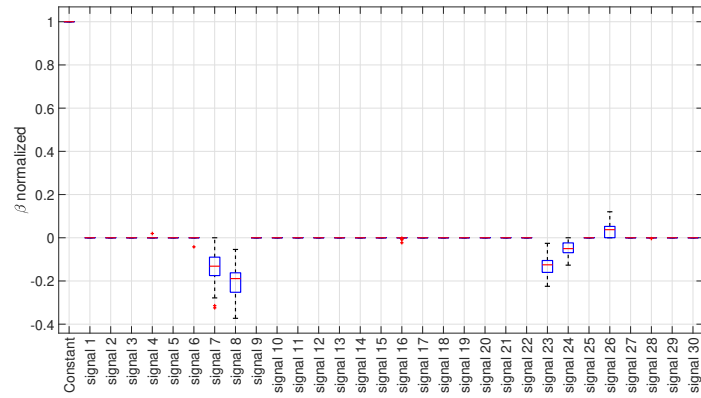
close to 3% which means that in the last part of the drying cycle (the focus for the automatic EoC) the estimation of laundry moisture content shows a median term that is close to the threshold allowed for the target (3% for Cotton). Each boxplot is the distribution of 100 iterations of the Linear Regression using a different set of training and test observations (details about the size of each set are reported in the next paragraph). Figure 10.1b shows the normalized β values obtained for MLR from the CV iterations according to the equation represented in Figure 4.1 in terms of distribution. In particular they refer to the first configuration depicted in Figure 10.1a (LASSO1) imposing the use of a maximum number of predictors equal to 5 for each iteration. Some observations are reported here to sum up the main results achieved with the approach at hand:

1. The solution is simple and without particular effort for implementation: the use of a linear models makes the solution very easy because the signals used as inputs are already available online on WD-HP machines and the information required is restricted to the values of β coefficients obtained with offline simulations;
2. With a small set of input signals is possible to achieve results with a RMSE error expressed with the same unit of the target measure (percentage) between 3% and 4% which is not sufficient to satisfy the target required by the industrial partner, but it is an interesting result considering the simplicity of the method;
3. Using a bigger set of input signals leads to an improvement in terms of performance and it seems not to be not possible to obtain moisture estimates with a RMSE of under 2% for the EoC using this kind of approach (Multiple Linear Regression using the sensor outputs as regressors) and the set of signals available.

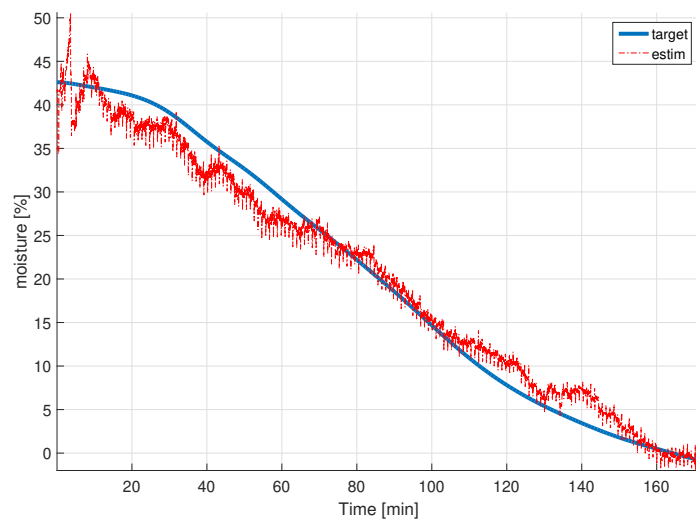
In Figure 10.1c the moisture target for one of the available observations together with the related estimation is depicted. Following this method, an update for the estimated value for laundry moisture is provided for each time sample ($[s]$). As the reader can easily see, the estimation is not smooth and this is due to the fact that signals used as predictors are affected by noise. This method is the simplest one proposed here because the moisture is estimated using signals already available as predictors combined in a linear model. However, it is considered not suitable for other problems of interest such as the TTE (Time-to-End) estimation problem which consists of determining how much time is needed to reach the laundry moisture target. It is not possible to estimate the TTE a long time in advance during the drying cycle using the method at hand because of the estimated moisture for a specific time sample changes according to the value assumed by each predictor at the time sample itself.



(a)



(b)



(c)

Figure 10.1: Examples of results using MLR: (a) performance in terms of RMSE distribution (in [%] of moisture content), increasing size of input signals from left to right; (b) variable selection made by the LASSO imposing the use of few predictors (LASSO1) for each iteration. The term Constant on x-axis is the intercept of the linear models for which coefficients are represented here; (c) example of laundry moisture content estimation using all predictors (RR), on y-axis the moisture content as a percentage.

10.2 Symbolic Regression

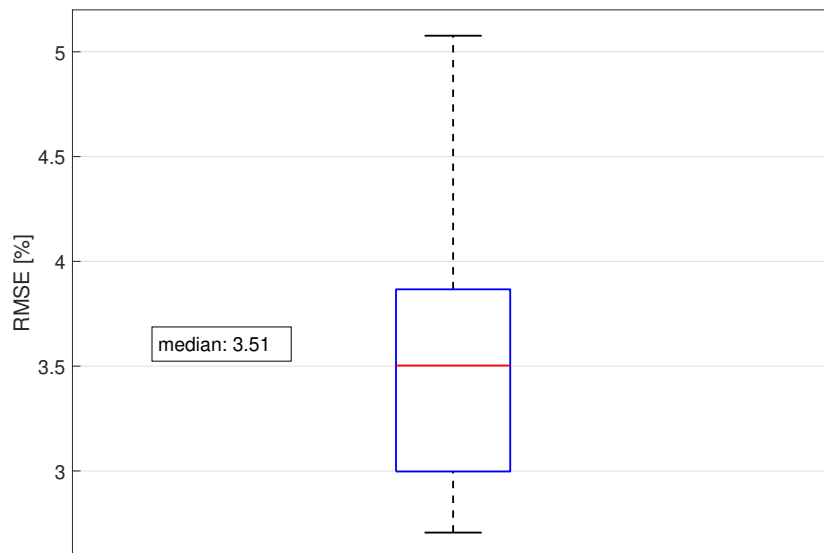
Figure 10.2a refers to the Symbolic Regression with GP approach and shows results obtained in terms of moisture error taking into account the final part of the drying cycle which is the main part of interest for an automatic procedure for the EoC.

The entire set of available signals have been exploited here to collect the results. The distribution of the obtained solutions is visualized in Figure 10.2a. Each iteration can give a different result in terms of the structure of the selected model, e.g. one solution is the following:

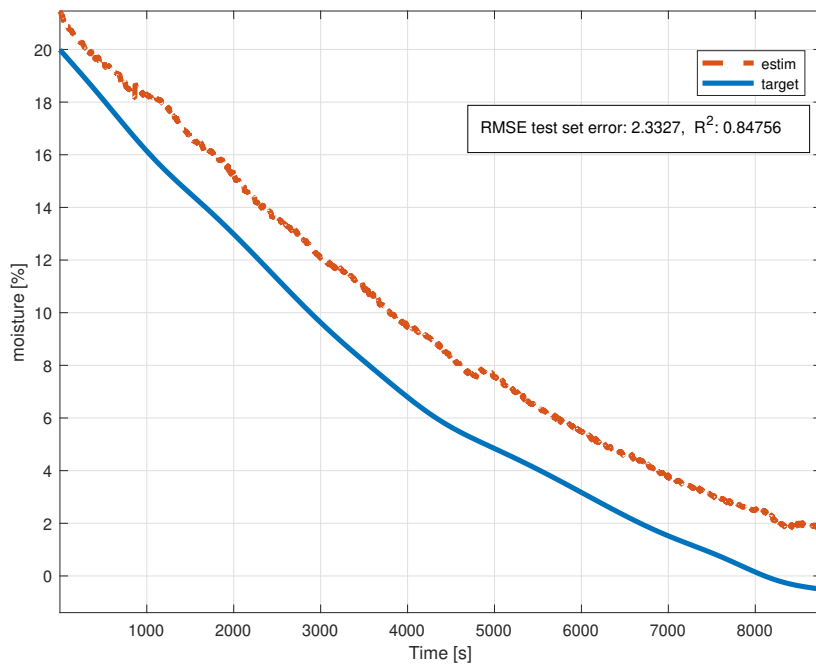
$$\begin{aligned}\hat{\mathbf{y}} = & k_0 + k_1 \mathbf{x}_2^3 + k_2 \mathbf{x}_7 \mathbf{x}_2 + k_3 \mathbf{x}_3 + k_4 \mathbf{x}_7 + k_5 \mathbf{x}_8 \\ & + k_6 \mathbf{x}_{16} + k_7 \mathbf{x}_{21} + k_8 \mathbf{x}_{23} + k_9 \mathbf{x}_{26} + k_{10} \mathbf{x}_{27}\end{aligned}\quad (10.1)$$

where $k_i, i = 1, \dots, 10$ are model coefficients and $\mathbf{x}_i$ are signals.

The solution selected at the next iteration could show different powers of the involved input signals, but the goal here is the discovery of the most useful models for the purpose i.e. which structures have been selected more frequently across all iterations (if any). It turns out that the most selected structures are the ones which involve the main temperatures and motor torque signals (or signals that are related to those); this confirms that these are also the meaningful signals for the goal. Furthermore, polynomial terms are selected mostly (instead of exponentials or trigonometric functions). Results from MLR and SR suggest that non-linear combinations of input signals do not improve the performance in general and it is supposed to be due to the problem of overfitting which is still present even though much reduced.



(a)



(b)

Figure 10.2: Examples of results using Symbolic Regression: (a) performance in terms of RMSE distribution (in [%] of moisture content) using Symbolic Regression with GP approach and all input signals available; (b) example of laundry moisture content estimation, on y-axis the moisture content in percentage.

10.3 Polynomial model

In Figure 10.3a results in the form of error distributions are represented. In this case the variability in the error is high and this is due to the variability of the estimation of the parameters of the polynomial model. Nevertheless, the focus now is on the reduction of this variability. In particular, Figure 10.3a shows the results for each laundry load amount considering the fact that the algorithm exploits a load estimation phase (as represented in Figure 9.1).

As depicted in Figure 10.3a median of each distribution (in percentage of moisture) is very low and always under the limit of 3% which is considered the threshold for an acceptable performance for the End-of-the cycle determination for the cotton laundry typology which is the case study at hand. However, the current problem which is still present is the error variability as explained before. Figure 10.3b shows also an example of estimation of the laundry moisture content during drying made using the polynomial model compared with the reference for the specific observation. It is noted that the use of this approach (polynomial shape of the moisture estimation) is preferred as a solution because of its smoothness. In fact in this case, inputs are features computed from signals and used to estimate parameters of the fixed model. Once the estimation of parameters of the model is ready, the result is a 3^{rd} degree polynomial in time which is smooth in contrast to results obtained with other approaches.

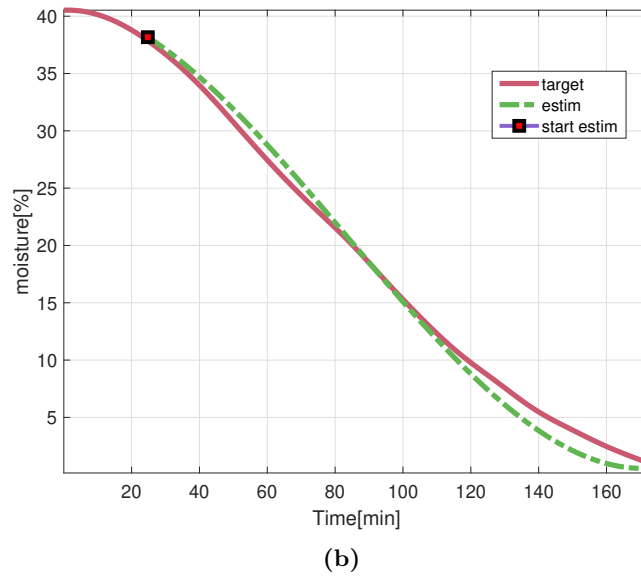
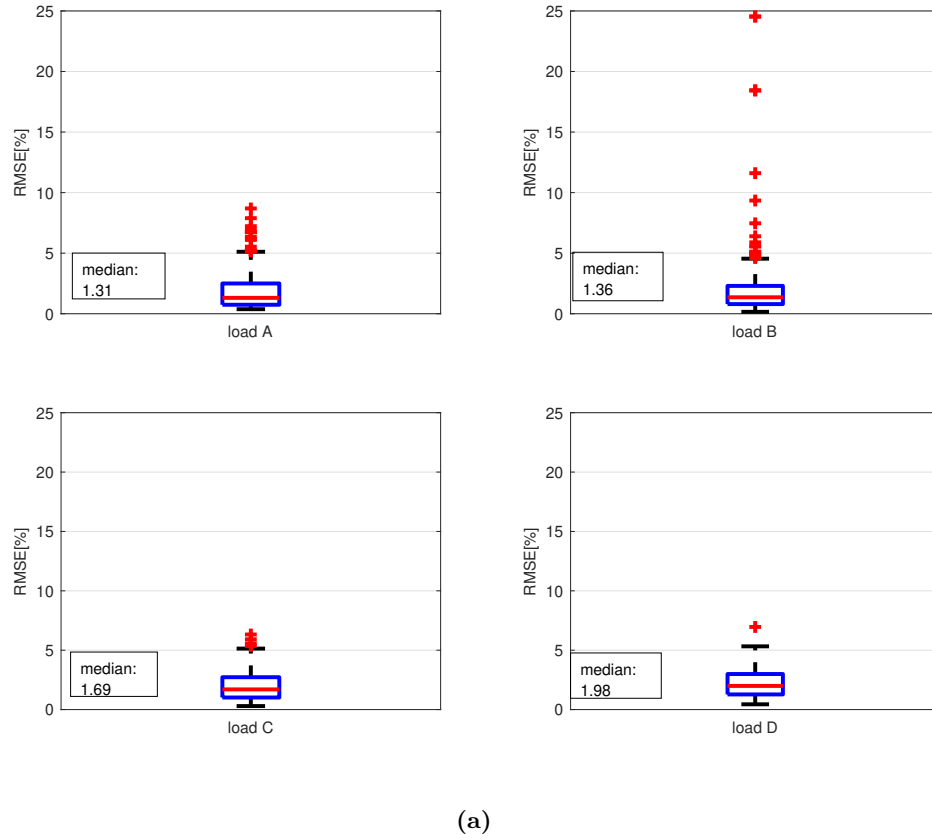


Figure 10.3: Examples of results using the polynomial SS (poly3): **(a)** Current performance in terms of RMSE distribution (in [%] of moisture content). Results divided by laundry nominal load; **(b)** An example of laundry moisture content estimation using the 3rd order polynomial model: on the y-axis the moisture content as a percentage [the square represents the time at which the estimation is provided during drying cycle].

10.4 Comparison between methods

Most of the research treated in this dissertation has been conducted on the design of SSs for laundry moisture content estimation during a drying cycles in the machines of interest. In order to organize ideas and especially to provide research answers to our initial question related to the feasibility of SSs in fabric care drying technologies, some remarks are formalized here below.

Methods	EoC feasibility	TTE feasibility	performance target*	low error variance	algorithm calibration	online facility
MLR	✓			✓	✓	✓
SR	✓			✓	✓	
poly3	✓	✓	✓		✓	✓
Current algorithm	✓	✓	✓	✓		✓

Table 10.1: Comparison between the proposed methods; * = the target for performance refers to the EoC estimation error, i.e. the error between the moisture reference and the estimate. The target formalized is to remain under a certain value of the error quantified using the RMSE metric defined before.

Table 10.1 summarizes the comparison between (data-driven) methods proposed in this work as procedures to estimate the laundry moisture evolution on household drying machines during drying process. The method named Current algorithm¹ refers to the current procedure implemented online on machines and it is mentioned here in order to make clear the features of each proposed procedure and the contribution for each item of interest².

As regards the criteria mentioned in Table 10.1, EoC and TTE feasibility refer to the possibility of carrying out the automatic End-of-the-Cycle and Time-to-End using one method or another. In particular, the interest is the chance to compute at the beginning of the cycle how much time is needed to reach the estimated moisture target and, in this respect, methods that provide an estimation of moisture which depends on the current values of the input signals (like MLR and SR), are not suitable for a direct computation of the TTE. In these cases the automatic EoC is possible because it is determined when the (estimated) moisture target is reached, on the other hand, the TTE task requires in these cases an additional dedicated algorithm. So, it is not possible to use the provided estimation directly to determine the TTE. In this perspective, poly3 is quite different because it allows an estimation of the laundry moisture evolution as a time series using

¹Details about it will be omitted here because of intellectual property rights.

²Items summarized here are the main requests made by the industrial partner with the goal of selecting a new algorithm for the control of the drying cycles on machines. The main focus is clearly on performance but also on the complexity and calibration effort in order to avoid some time-consuming procedures involved in the use of the current algorithm.

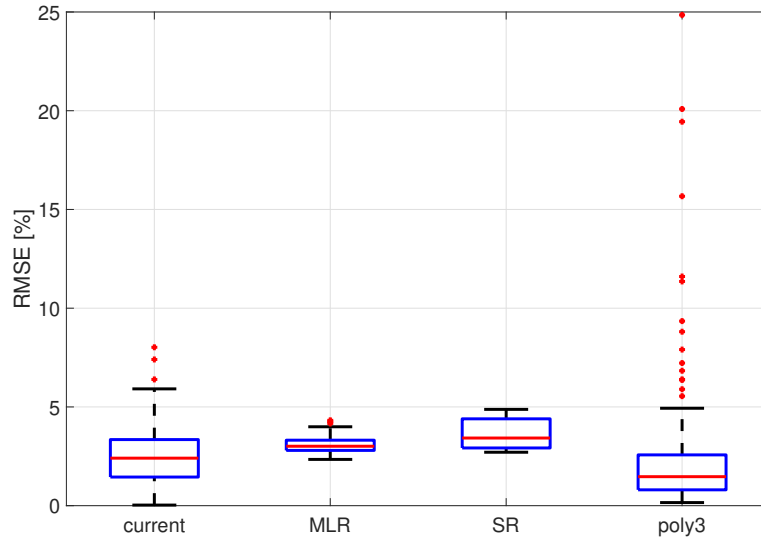


Figure 10.4: Laundry moisture SSs: comparison between methods in terms of estimation error distribution.

features computed at the beginning of the drying cycle. It permits an evaluation of the TTE once the model is available (after 20 minutes from the start phase). In fact, once the estimation of the parameters of the poly3 is available (equation (9.1)), it is possible to determine the time needed to reach the moisture target y .

Figures 10.1a and 10.2a show that MLR and SR in terms of performance are not suitable for the predefined target (error distribution under 3%), although acceptable in terms of variance. As seen in Figure 10.3a there is still a problem related to the variance of the estimation error in the use of poly3, especially for load B.

Some new constraints on estimated values of parameters of the cubic model (9.1) have been imposed in order to reduce the error variance taking into account the fact that drying tests in laboratory are time-consuming and no more data are available for this work now. The use of a richer dataset could be a next step of this project; here the goal was to develop SSs using already available tests made for different projects, avoiding dedicated Design of Experiments (DOE) tests. The last two columns in Table 10.1 show a comparison in terms of calibration and implementation effort. As regards the former, it turns out that the current algorithm in production suffers from a time-consuming calibration phase because operators have to fix some thresholds on the EoC signal and these thresholds have to be managed on firmware. In the latter case, SR suffers from an additional complexity level because models selected by Symbolic Regression could

include some nonlinear terms which are difficult to write on firmware, e.g., combinations of exponential, trigonometric functions, etc. The other methods are simpler to represent online because signals are combined in a linear fashion (MLR) or because the model is a time series with few parameters (poly3).

A comparison of the methodologies and procedures exploited in this part of the research is visualized in Figure 10.4 in terms of error distributions for the EoC goal. The current implemented procedure is compared to the other techniques (proposed SSs) and experimental conditions are the same. Two considerations seem to be arising from this kind of visualization:

- (i) there's no one solution that can be considered absolutely better than the current approach because the median or the variance of the distributions is worse. Nevertheless, SS called poly3 shows errors with a significant density under the level of the median error of the current approach, so, the priority to achieve definitive improvements is to limit the variance on results as much as possible;
- (ii) the MLR approach shows very good results in terms of variability with respect to the other cases and this behavior is partially expected due to the fact that the structure of the model here is fixed and results depend only on the trend of the input signals which are combined linearly. This practice helps to mitigate the overfitting problem which is the main enemy of the methods described in this part of the thesis.

Ridge Regression represented in Figure 10.1a was chosen to represent the MLR methodology here; a first reason is the fact that RR showed good performance with respect to the other regularization choice; a second reason is that we need to use the same (or very similar) experimental settings and, also considering SR, the total amount of the available signals (after preprocessing) was exploited. Focusing now a little bit more on poly3 results, the boxplot for poly3 in Figure 10.4 includes the performances considered for all the laundry loads studied from the dataset. It also highlights how the variance problem on results is still significant. This is a consequence of the difficulty in using powerful features (summarized in Table 9.2) when few data are available. In relation to this subject, it is important to consider that poly3 shows more variability on results when applied to small and medium laundry loads (load A and B in Figure 10.3a). This may be due to the lack of available data which is problematic for this method; to support this statement, a study on regression diagnostics was conducted as suggested in the methodological part of the thesis (4.1) dealing with the use and reliability of linear models. One result of this study is presented in Figure 10.5 where Q-Q plots appear.

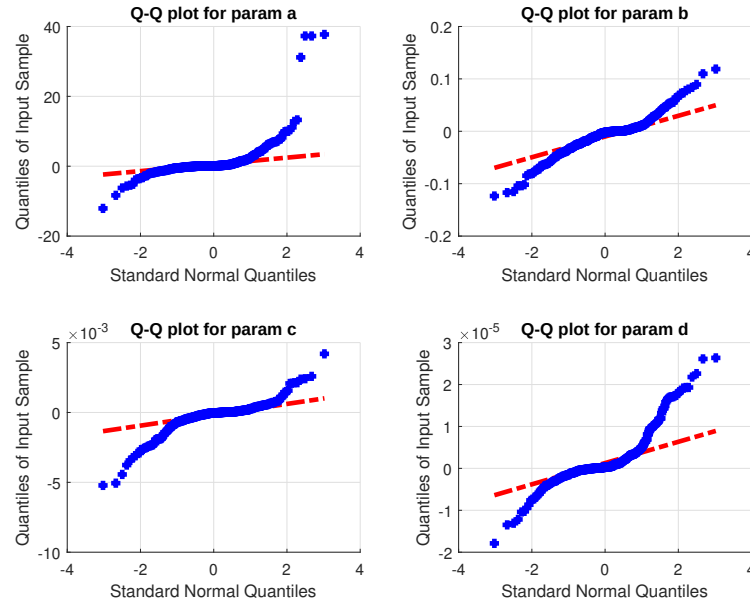
They are used in general to test if data come from a certain distribution (such as the Normal one) and are used here to test the normality of the residuals as a key assumption in the use of linear models. When represented samples deviated from the line it means that it makes sense to reject the hypothesis of normality, i.e., in these cases the use of linear models for the purpose becomes ineffective. In these kinds of plots, two sets of quantiles are shown against one another and if both sets of quantiles came from the same distribution, the points form a line that's roughly straight. So, in this case, samples from normal distributions are compared with the residuals from regression to produce the Q-Q plots. From this, observing Figure 10.5 it seems that the use of linear models for the estimations with data from load B is more difficult from the case of load D because, in D, Q-Q plots show a better behavior; it has indeed already been seen in Figure 10.3a that the best performance is achieved for load D data for which much more cycles are available in the dataset. So, considering this and the fact that we have few data for load A and B, we could justify the occasionally bad performance of poly3 SS as a consequence of ineffective use of linear models for the estimation of the parameters of the polynomial with the limited amount of available data. Laundry moisture signals related to certain laundry loads (A and B) are clearly hard to predict properly with few available data, so, some simple controls on the predicted curve have been adopted in poly3 SS (such as the one discussed in 9.3 based on the computation of the drying rate). Poly3 lends itself to this kind of controls more than the other proposed SSs because of the smoothness of the estimated curve.

As a final consideration, the application of several methods for the problem at hand has provided some achievement on the methodological point of view:

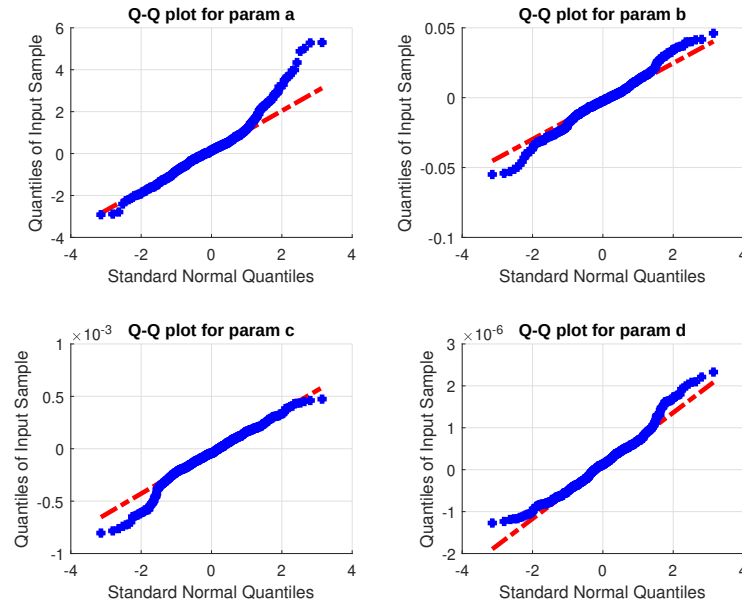
- the use of (regularized) linear models shows a limit in terms of estimation performance on the available dataset (boxplot medians in Figure 10.1a). Nevertheless, the use of more complex models in similar experimental settings does not provide better results; this surprising fact is likely due to the difficulty in limiting the overfitting in SR with the available data;
- MLR and SR methods have offered additional knowledge of the process enabling us to perform a more efficient feature engineering phase for the last procedure (poly3) being conscious of the key role of some signals among the others in the construction of a SS for the purpose;
- the inclusion of some steps in the poly3 algorithm allowed us to limit the variability observed testing the model offline: the sub-procedures called control on shape and dry laundry load estimation have actually this role as explained. Variability on

results was the main problem also with the other methodologies but in those cases, the use of these simple practices turned out to be more difficult since rough noisy signals were used as inputs in those cases. Furthermore, the poly3 approach can be used as a solution for the EoC goal (limiting the variability on results for certain laundry load amounts).

The lack of data is a problem for all the SSs designed in this part of the research but the results showed by poly3 in terms of the median of the error distributions are promising as long as the variability is handled properly, i.e., with the use of simple controls on the shape of the estimated curve.



(a) Q-Q plots for the estimation of poly3 parameters with data from of load B.



(b) Q-Q plots for the estimation of poly3 parameters with data from of load D.

Figure 10.5: Q-Q plots.

Part IV

Soft Sensors for laundry weight estimation

11

Dataset description

11.1 Heat Pump Tumble Dryer data

This part of the thesis is dedicated to the second problem of interest as a case study for SS design: the laundry weight estimation task. This topic has already been introduced in 3.3 dealing with the main algorithms studied for fabric care R&D drying technology applications and objects of continuous improvement. Since the core part of the dissertation (SS for laundry moisture estimation) has already been developed, this section will be much shorter. The data used for this section comes from a different machine; whereas in the previous case we were dealing with a WD-HP here the machine is a (Heat Pump) Tumble Dryer (TD-HP). There is no substantial difference in the process because in both cases we are talking about data acquired during drying cycles.

The main modules of the machines are still valid as well because both of them exploit the heat-pump technology (the main modules of the process have already been explained in 1.3 and 1.4). The dataset of interest is a part of a bigger set planned for the development of a new on machine physical moisture sensor (details will be omitted because of intellectual property rights); it is recalled here that TD-HP, differently from WD-HP machines, are equipped with physical moisture sensors, so in this case, there's no need for a laundry moisture SS but the focus is on the improvement of the current laundry weight estimation algorithm which exploits the measures acquired from the new physical moisture sensor.

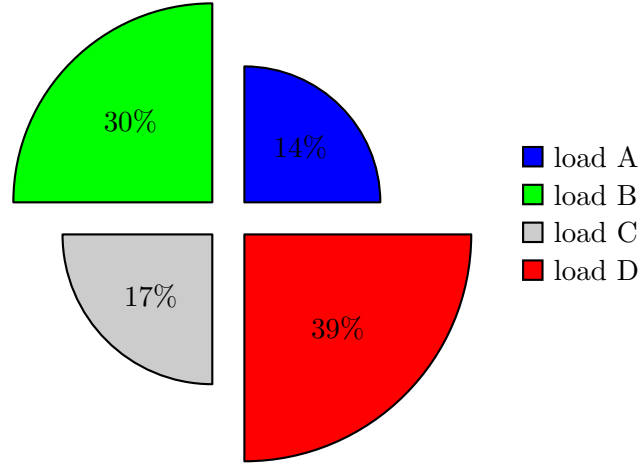


Figure 11.1: Tests in the dataset divided by nominal weight.



Figure 11.2: The indication of drying remaining time (TTE) on a dryer machine.

Such desired improvements are aimed to provide a more precise value of the TTE (discussed in 3.2) on the user interface through a conversion from estimated laundry weight and remaining drying time. An example of dryer machine is given in Figure 11.2 where it is possible to see also the user interface on which the TTE information appears after the first minutes of the cycle.

Lots of details about the dataset and the source are omitted here and the reason is that the dataset belongs to an on-going project, so lots of aspects are not to be disclosed. The available data distribution in terms of nominal weights (which are named here with the same fictitious labels used in the other case study but which are actually different) is given in Figure 11.1. As can be seen, the dataset is unbalanced with respect to the large loads (load D). Moreover, in this case, the dataset is richer: a total number of $n = 446$ entire drying cycles are available. As said above, these data come from several activities planned for an on-going project, hence the availability of much more data in this case

study. It has been stated in this dissertation that such a situation is not common for in our context, i.e., generally, only tens of entire drying cycles are accessible to perform evaluations. So in this particular situation, there is the possibility to compare the current algorithm implemented online (already introduced in 3.3) with a different technique using hundreds of tests.

The remaining part of this section will be dedicated to the explanation of the new approach proposed to address the problem of the design of a SS for laundry weight estimation in order to compare it with the current implemented procedure presented in 3.3. This dataset was chosen among different others because of the current interest in laundry weight estimation improvement related to the new physical sensor. So it becomes interesting to test it with some different techniques such as the functional regression approaches introduced in 4.3. To the best of my knowledge, this is the first application of a functional regression technique in the context of real data coming from the development of fabric care R&D projects.

12

Scalar-on-Function Regression models

12.1 Description of the procedure: Functional Linear Models (FLM)

This kind of approach was introduced in 4.3 dealing with Penalized Functional Regression Methods. Some other details are covered here focusing on the Scalar-on-Function regression problem because it matches our requirements, i.e., the estimation of a scalar value (the laundry nominal weight) using entire signals (functions of time) as predictors. It is pointed out that the goal in this case, is the final estimation of a scalar quantity (laundry weight) and not a time series (laundry moisture) as in the previous part of the thesis. So, just to make it clear; in this part of the research we have:

- *inputs*: entire signals from real acquisitions on TD machines equipped with the new physical moisture sensor;
- *output*: β -functions and laundry dry weight estimation obtained offline from Scalar-on-Function regression using signals as predictors;
- *reference for model training*: the reference used as the target is the laundry nominal weight (laundry without water) available from laboratory tests.

The use of the functional approach described in this part of the thesis seems to be justified in terms of the goals we want to achieve to make clear in which sense a functional procedure could be useful. In this part of the research, the objective is not only the consideration of estimation performance (compared to the one obtained with MLR and explained in 3.3), but the main purpose is to obtain representations of β functions in order to understand the role of predictors in estimation over time. These kinds of representations could improve the current feature engineering process performed today with MLR: in fact, the knowledge of the precise contribution of a given predictor to the estimate lets us be more selective in the process avoiding the use of some signals or considering only some time intervals of a certain set of predictors. This concept appears more vivid observing Figure 12.1 where the β -function has been framed to stress which is our focus.

Differently from what has been seen so far with MLR, β s are here no longer values but functions, and therefore their estimations also provides a feedback on how each predictor contributes to the estimation in time: the more the absolute value of the estimated beta function deviates from zero in a specific time interval, the more the related predictor affects the estimation.

Going now into details of the methodology we adopt, from a practical point of view, the first question is why we need functional techniques in this case and in which sense data are said to be functional. Now, according to observations made by Noguerales (2015), it is actually impossible to observe complete functional data. Nowadays, high-tech sensors allow for monitoring processes on finer and finer grids, but in the end some sort of discretization is used.

Therefore, what we have in practice is not the process $\{X(t), t \in [0, T]\}$ but a high-dimensional vector $(X(t_1), \dots, X(t_N))$ in the discrete grid $t_1, \dots, t_N$. Therefore, a question to discuss is if we can talk about true functional data. In my opinion the answer is affirmative: on the one hand, in some cases, the grid can be as fine as desired so the process is virtually observable at any instant t . On the other hand, we can choose a functional model to approximate the data. So, the idea is that behind the vectors we have there are real functions with all the corresponding implications, or in the words of Ramsay and Silverman (2005): “The term functional in reference to observed data refers to the intrinsic structure of the data rather than to their explicit form”.

Furthermore, the presence of noise presents similar problems. In practice, most functional data are contaminated with random noise. This is also called measurement error, although sometimes the source of noise will not be an error in the measurement. Noise is usually modelled considering that we observe a data $y(t) = X(t) + \epsilon(t)$, where $X(t)$ is a real

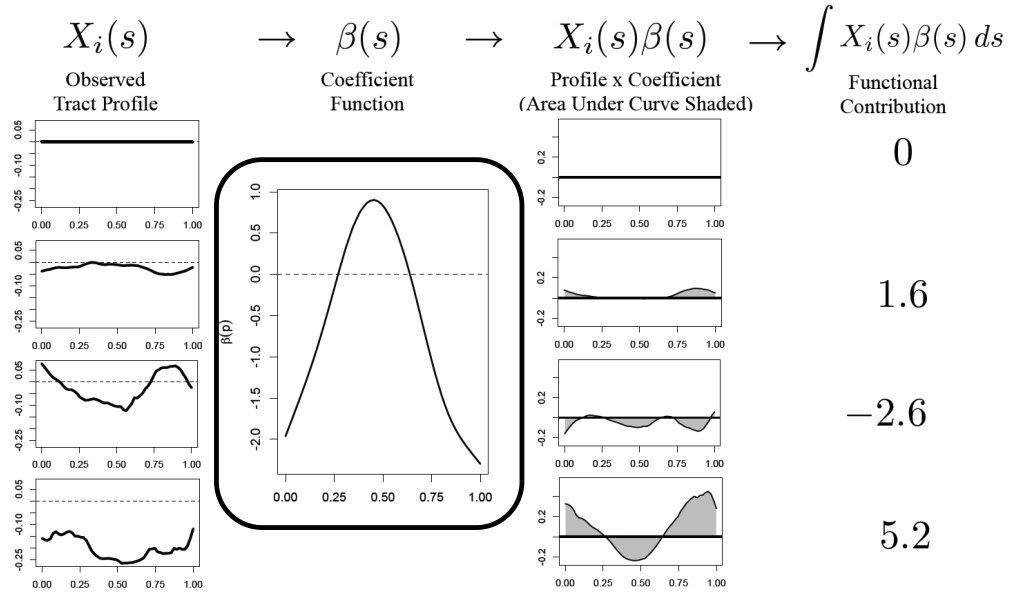


Figure 12.1: Interpretation of β -functions in FLM. Taken from [Ogden and Goldsmith \(2014\)](#).

function and $\epsilon(t)$ is random noise independent from $X(t)$ such that $E[\epsilon(t)] = 0$ and $Var\{\epsilon(t)\} = \sigma^2$. These errors can be sometimes insignificant, but in other cases, the noise is a critical point, here in our case study, for example, noisy signals are used to predict a scalar quantity whose precision is significant for the cycle handling. The goal in both cases is obtaining the original function from the (noisy) observed vector, or sometimes, getting a smoother version for further actions, for example, to take derivatives. In [4.3](#), FDA has been introduced as a set of techniques to perform functional studies on data and relative predictions.

There are FDA tools that allow us to “recover” the original function (or a very close approximation) from the observation or just get a smooth enough approximation. These methods are basically grouped into two big families: basis representation and smoothing. The basis representation is a recurrent tool in FDA. It plays the double role of getting a continuous function and reducing the dimension by truncating the expansion series. A full review of these topics is not covered here but it can be found in [Ramsay and Silverman \(2005\)](#)-chapter 3. Assuming that the data $X(t)$ live in a functional space $\mathcal{F}$ and let $\{\phi_i(t)\}_N$ be a basis of that space, then $X(t)$ can be represented by the series $X = \sum_{i=1}^{\infty} c_i \phi_i(t)$, where c_i are the coefficients corresponding to the basis. In practice, the infinite series is truncated at certain element k and we approximate $X(t)$ by $\tilde{X}(t) = \sum_{i=1}^k c_i \phi_i(t)$. This number k is a sort of smoothing parameter that must be carefully chosen. Larger values of k produce good approximations (perhaps leading to

over-fitting) but keep the high dimensionality problem. Besides, smaller values of k make the data easy to handle but some important information could be lost.

Unluckily, there is no general rule to set the best number of elements and a part of the research described here was dedicated to the discovery of suitable values of this number in our case study as will be stated in the following. The functional space $\mathcal{F}$ is typically supposed to be $L^2[0, 1]$. In this case there exist an orthonormal basis $\{e_i\}$ and the coefficients can be easily calculated by means of the inner product, $\tilde{X}(t) = \sum_{i=1}^k \langle X(t), e_i(t) \rangle e_i(y)$. Of course, the properties of $\tilde{X}(t)$ depend on the basis functions. Hence, an adequate choice of the basis, according to the nature of the data, is needed.

Among the wide variety of functional basis (exponential, polynomial, etc.) maybe the most frequently used is the Fourier basis since it is very easy to derive and the coefficients can be efficiently calculated via the Fast Fourier Transform algorithm (FFT). Splines basis might be the most popular basis nowadays (for non-periodical data) and it is also the choice made in the research at hand here. This basis combines the efficiency of polynomials (which are included in it) with greater flexibility so that it usually needs just a few elements. The technique basically lies in dividing the time interval and making a polynomial approximation in each subinterval while taking care of the breakpoints. No more details are covered here about the theory of splines; a reference suggested to the interested reader is for example, [Ahlberg, Nilson, and Walsh \(1967\)](#) which has been taken into account in the development of the research at hand.

The main steps of Scalar-on-Function regression as a Penalized Functional method, or, more properly, Functional Linear Models (FLM) method, have been summarized here for reasons of clarity. The reader can assume $X_i(s)$ as an input signal (variable s used for time) and Y_i the scalar output value (laundry nominal weight) as usual.

- **Step 1:** estimation the random functions X_i using a finite series expansion and use of truncated power series spline basis to represent β :

$$\begin{aligned} X_i(s) &= \sum_{k=1}^{K_x} c_{ik} \psi_k(s) \\ \beta(s) &= \sum_{k=1}^{K_b} \theta_k \psi_k(s) \end{aligned} \tag{12.1}$$

in red the two critical choices for this method as stated before.

- **Step 2:** perform the output basis function representation starting from equation seen in (4.2) and exploiting the assumption of linearity which is still valid in this

functional approach.

$$\begin{aligned}
Y_i &= \alpha + \int X_i(s)\beta(s)ds + \epsilon_i \\
&= \alpha + \int \left(\sum_{l=1}^{K_x} c_{il}\psi_l(s) \right) \left(\sum_{k=1}^{K_b} \theta_k\psi_k(s) \right) ds + \epsilon_i \\
&= \alpha + \sum_{k=1}^{K_b} \left[\sum_{l=1}^{K_x} c_{il} \left(\int \psi_l(s)\psi_k(s)ds \right) \right] \theta_k + \epsilon_i \\
&= \sum_{k=1}^{K_b} z_k\theta_k + \epsilon_i
\end{aligned} \tag{12.2}$$

- **Step 3:** perform the regularization. A first term is used to measure fidelity to the data (how well the model “fits” the data, while a second term measures the smoothness of the coefficient function.

$$\sum_{i=1}^n \left(Y_i - \alpha - \int X_i(s)\beta(s)dt \right)^2 + \lambda \int (\beta''(s))^2 dt \tag{12.3}$$

$$CV(\lambda) = \sum_{i=1}^n \left(Y_i - \alpha_{\lambda}^{(i)} - \int X_i(t)\beta_{\lambda}^{(i)}(t)dt \right)^2 \tag{12.4}$$

λ value is chosen with Cross-Validation minimizing the expression called $CV(\lambda)$.

The choice made to compute these steps will be justified in 13.1. Only a remark is provided here because of its importance for the use of the methodology at hand: the truncated power series representation of $\beta(t)$ imposes differentiability and allows simple control of smoothness; this is important for the use of the proposed regularization.

13

Experimental results

13.1 Experimental settings

Implementation details of the Scalar-on-Function regression method are summarized in this part together with the main results obtained. The experimental environment employed is as follows:

- First of all, the entire analysis has been conducted without considering complete drying cycles: indeed, this part of the research is focused on the estimation of the laundry weight amount during the first cycle phase. As already explained in the previous sections of the thesis (3.2 and 3.3), the load estimation is used for the determination of the TTE (the initial computation of the TTE, because other updates are made during drying). So, also in this research activity, the focus has been given to the first part of the cycles limiting the observation to the available data in the first 20 minutes. For this reason, the results presented here always show a time reference that is $t = [0, 20]min$ and this means that the information related to the rest of the drying process has not been exploited in this analysis.

It is pointed out that the choice is not restrictive, i.e., a lower time interval could be used, but in this case the idea is to investigate possible improvements in using the first 20 minutes of acquisitions respect to the current performance achieved

with an algorithm that determines the dry weight in the first minutes of the cycle.

- As announced, the critical step of the procedure at hand is the choice of the K_x and K_b values by which the inputs and β are parametrized. It is suffice is to say that it is always a trade-off solution: K_x must be large enough that $\beta(t)$ is in the space spanned by $\psi_1(t), \dots, \psi_{K_x}(t)$, but small enough to smoothly estimate $\beta(t)$ and this is because we are using the same eigenfunctions ψ for both the predictors and β . From this, one factor that needs to be taken into account is an identifiability constraint: select $K_x \geq K_b$. Apart of this, no other criteria are a priori valid.

The decision made in this case study was to consider how this problem has been addressed in literature in similar problems (when noisy input signals are involved and the requirement of interpretability is claimed), as in, for example: [McLean et al. \(2014\)](#). In [McLean et al. \(2014\)](#), a choice for K_x and K_b beyond the value of 10 was not considered because for their problem it did not provide benefits in terms of better estimation performance. Some other works were taken into account, such as [Ruppert \(2002\)](#) and [Goldsmith, Bobb, Crainiceanu, Caffo, and Reich \(2011\)](#), in which some algorithms for the choice at hand is discussed. These algorithms follow a common line that is the execution of some CV statistics given a sequence of trials; this is also what has been done in the case at hand executing several simulations with different configurations of K_x, K_b (respecting the identifiability constraint).

A final trade-off solution was identified as $K_x = 10, K_b = 5$ for our case study: greater values for K_x do not provide substantial differences in performance and $K_b > 5$ does not give visualizations with different shapes for β -functions. The general idea learned from the literature and followed also in my research here is simply: make certain that K_x and K_b are sufficiently large to fit the data and not so large that computation time is unnecessarily large. It must be said that a more suitable method to achieve this knowledge could be the application of some multi-objective optimization algorithms, for example using the Genetic Programming paradigm discussed for Symbolic Regression in this dissertation (4.2). This direction could be taken into account to confirm/verify the decision we made.

- the starting point for the code developed in this part was the R `pfr` function in `refund` package ([ref \(2018\)](#)). The main steps discussed so far inclusive of the Cross-Validation algorithm (K-fold automatically optimized) for the selection of the regularization parameter are done using this function. In particular REML method is used for smoothing predictors using the decided value for k_x ([Oehlert \(2014\)](#)). REML is a way to estimate variance components. Once we have estimated variance components, we then assume that the estimated components are “correct” (that is,

equal to their estimated values) and compute generalized least squares estimates of the fixed effects parameters.

- the analysis proposed here has been done with a piece of certain previous knowledge about the most useful predictors for the problem. This a priori knowledge has been achieved in these years working in several R&D projects where laundry weight estimation was involved as a step to create the algorithm for TTE. That is the reason why this part of research has value: we already know which signals include the main information for our estimation goal, so we can use functional approaches to achieve 2 types of results: (i) the confirmation/verification about our hypotheses observing results from Penalized Functional Regression; (ii) the discovery of new time intervals (between 0 and 20 min) that are not considered yet for the improvement of the estimation process. The results summarized here below are discussed from these perspectives.
- each acquisition (drying cycle) consists of 8 signals.

The main results of this activity are visualized in Figure 13.1 where the estimation error ($\hat{Y} - Y$), $\hat{Y}$ for the estimation and Y for the real nominal weight) is represented compared with the one that comes from evaluations of the current implemented algorithm obtained from real tests in laboratory. In this kind of evaluation, all predictors have been exploited to obtain the estimation. It is recalled that, the main difference between the two methods is the fact that the current one uses features computed in the first minutes of the cycle (procedure described in details here: 3.3), while FLM method exploits the entire set of signals in the time interval $t = [0, 20]min$. In both of them, the assumption of linearity between the inputs and output is still valid.

Unlike the error visualizations used for the other problem (moisture estimation), for example, Figure 10.4, in this case, the estimation error is represented with as the difference (estimation-nominal weight) instead of the RMSE. This choice is motivated in this thesis as a specific request for the problem at hand: we need to know if the designed SS tends to overestimate (negative result and so the SS detects less laundry than the truth) or underestimate the target (positive result and so the SS detects more laundry than the truth). Therefore, also the sign of the error is significant in this case because positive errors are preferred for the design of this specific SS.

Results show that, a complete improvement in terms of estimation performance has not been achieved observing the variability of FLM, but slight improvements comes from the fact that the maximum absolute value of the FLM distribution is less than the one observed for the current solution; on the contrary, one negative aspect of the proposed

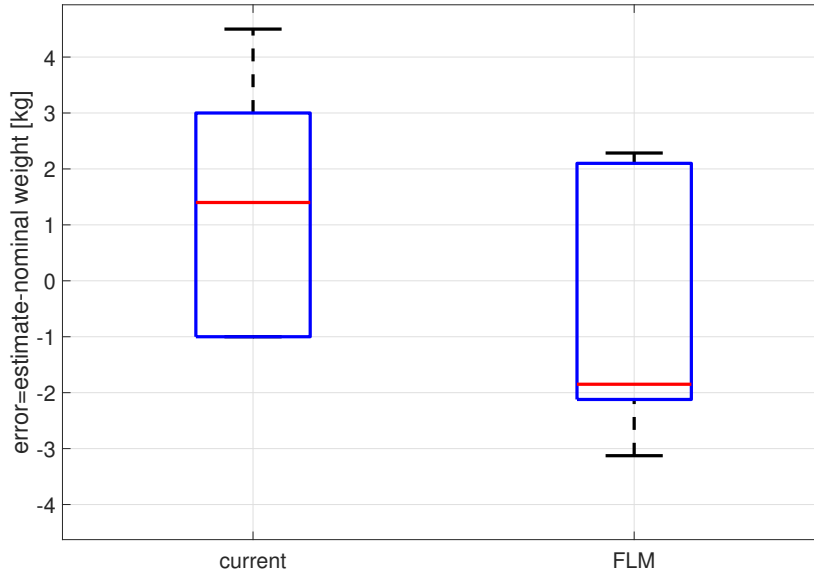


Figure 13.1: Laundry weight SSs: comparison between methods in terms of the error distribution.

solution is the fact that it tends to underestimate the laundry weight. In fact, it is better to see over-estimation trends because we would like to avoid situations where the user obtain partially wet laundry at the end of a drying cycle caused by controls based on such an estimation.

The achieved result is more clearly observing Figure 13.2 which shows the real goal of this research: the β -functions associated with the relative predictors (signals 1 to 8). In this kind of evaluation, each signal is used as the unique predictor to obtain the estimation. Two considerations are made here based on these results:

- (a) Some estimates seem to be unreliable because of the shape of the visualized confidence interval (signals 4,7 in the Figure);
- (b) the remaining set of signals shows a variable contribution to the estimates, in particular, at the beginning of the cycle and close to $t = 20[min]$.

In remarks to the above considerations, a study was conducted on regression diagnostics as done for the moisture estimation case (Figure 10.5). In this evaluation, each signal has been used as the unique predictor to obtain the estimates; after that, the associated residuals were studied. This study confirmed what was discussed in the previous (a) statement because Figure 13.3 underlines problems in Q-Q plots for the residuals (signal

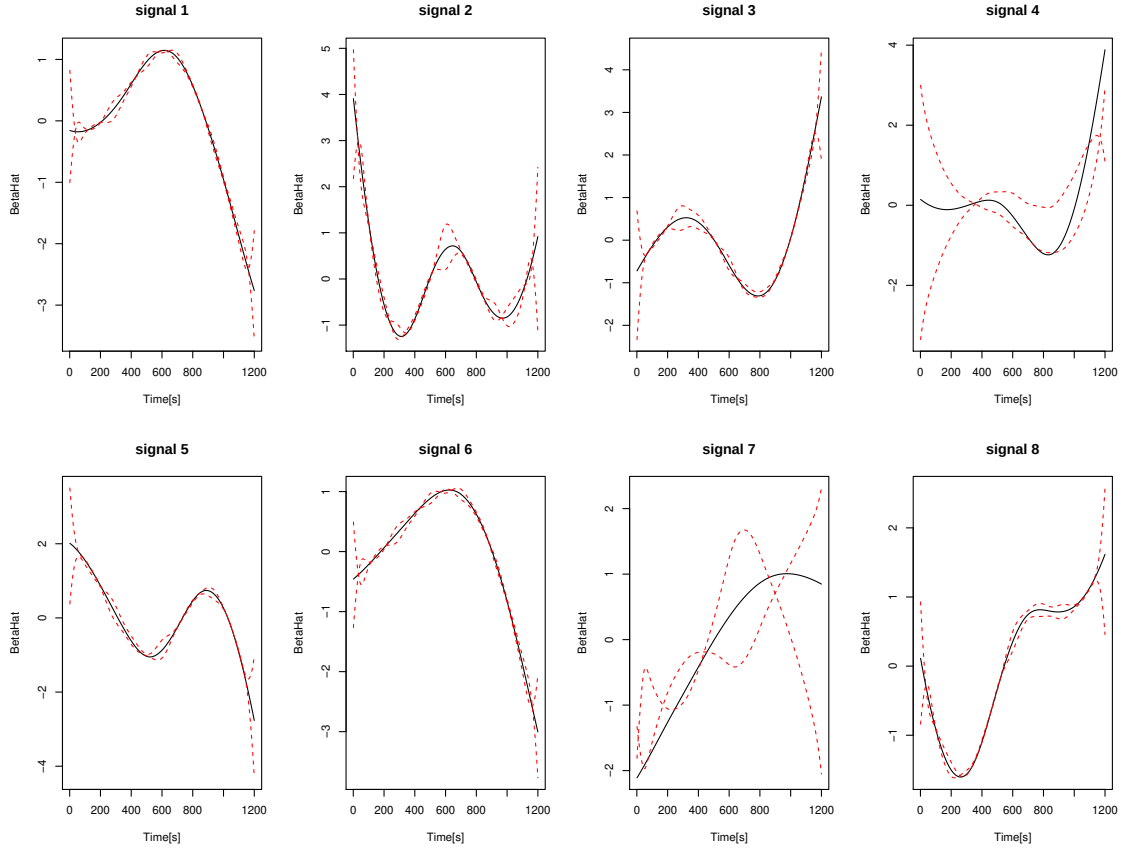


Figure 13.2: β functions obtained using FLM approach. Values on y-axis scaled in mean and standard deviation.

4, signal 7); therefore it is not suitable to make interpretations for β -functions related to signals 4 and 7. Other types of visualizations introduced in 4.1 were produced, such as residual plots or distribution plots but they are not reported here since they provide the same kind of conclusion about the problem found.

A special remark is made here in relation to signal 2. This signal is the one related to our a priori knowledge on the problem, i.e., we know that it has a significant role when it is used for laundry weight estimation. It had been actually computed also for this purpose. Observing the β -function for signal 2 one result is a confirmation of its role in the problem. In fact this β -function is the one that shows the higher absolute value in the first part of the cycle and this is exactly the reason why features from signal 2 are computed by the current implemented algorithm during the first minutes of the drying. This behavior is explained better in Figure 13.4 where signal 2 from 4 acquisitions of different laundry weight is depicted. The first part of drying cycle in which the signal

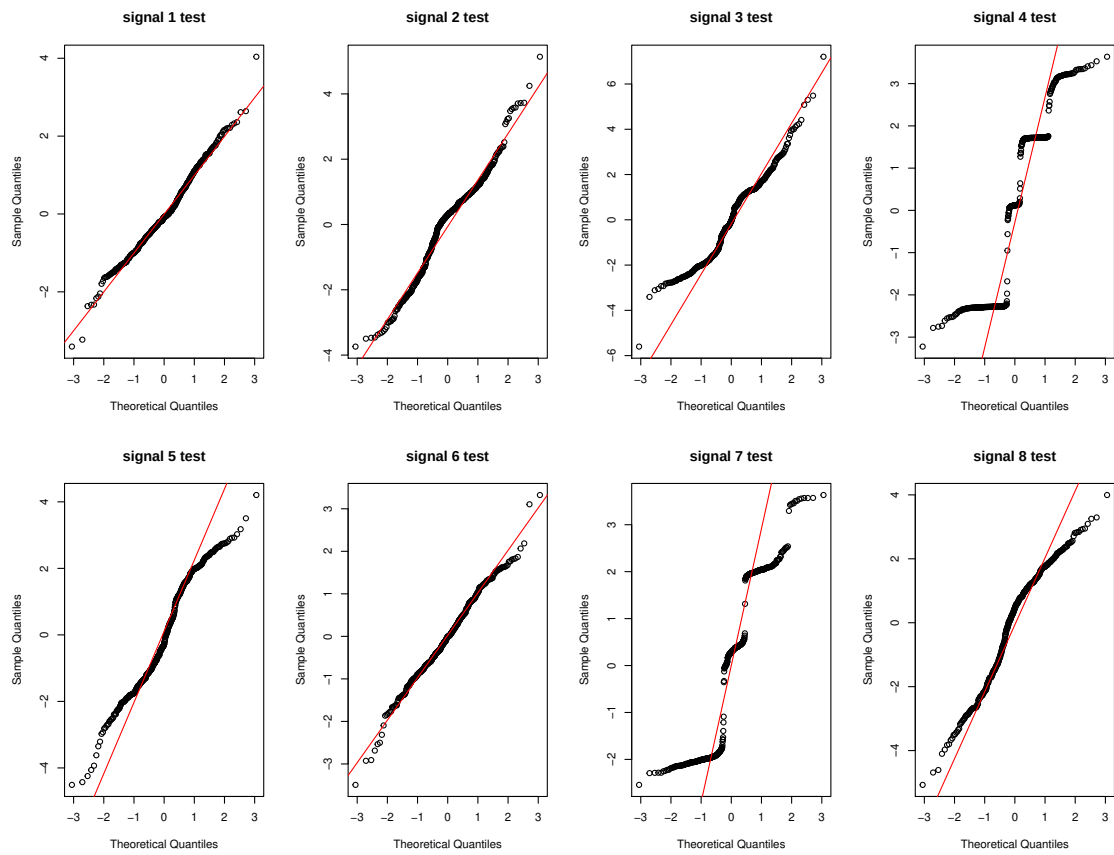
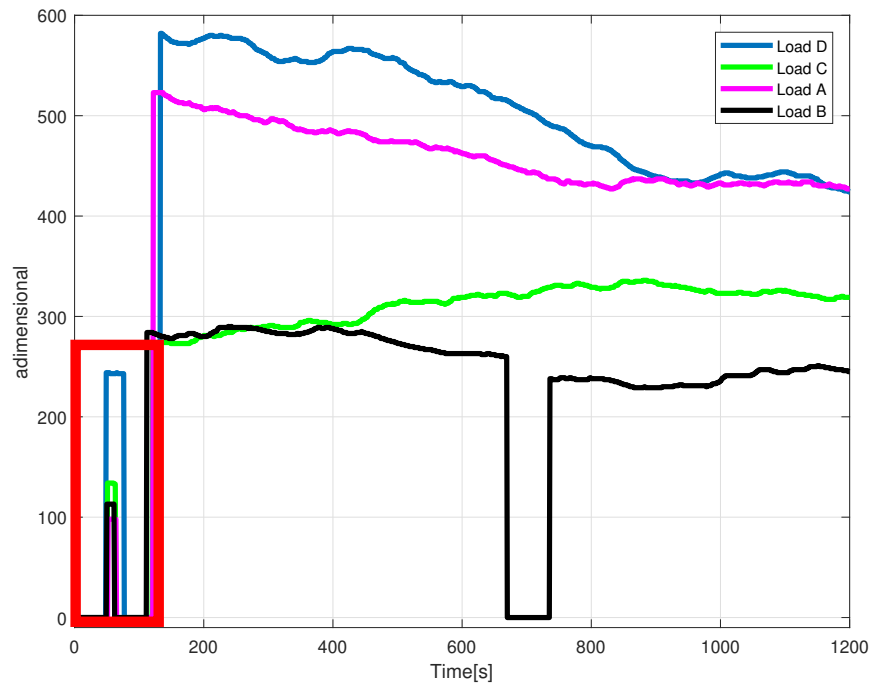


Figure 13.3: Q-Q plots associated to estimations made using each signal as unique predictor.

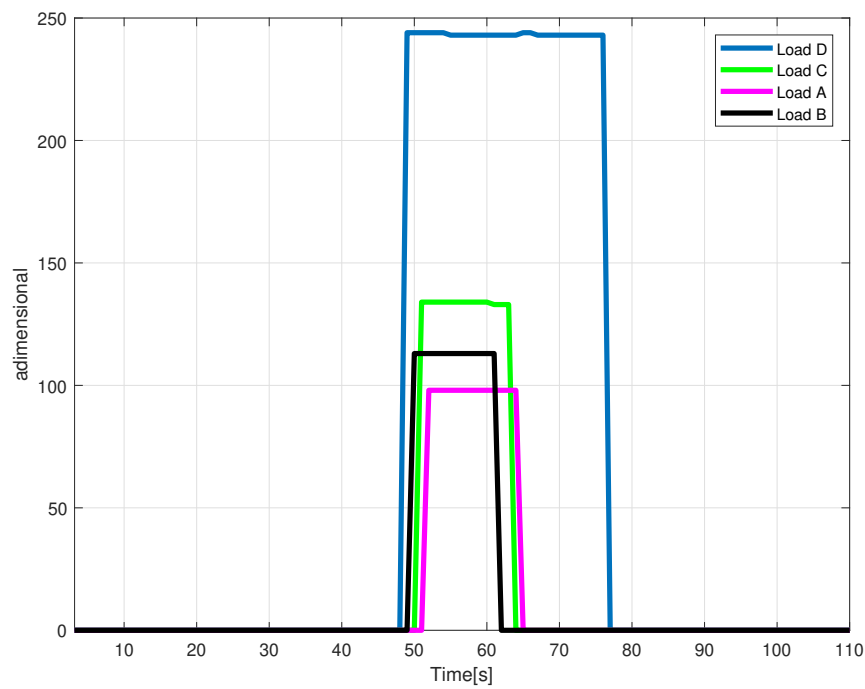
is used for feature extraction has been framed; it seems clear from the plot how the acquisitions are well-clustered in the first time interval. Apart from signal 2, in the other cases a significant contribution to the estimate is given close to the final part of the time interval considered (Figure 13.2 signals 1, 3, 6), so, for example no other suggestions come from the observation of the plots if we are interested in the improvement of the current approach using information from the first 5 minutes. As another observation, signals 1 and 6 are characterized by a similar shape of β -function, this is due to the fact that they are both temperature signals: one comes from an NTC sensor inside the machine while the other is embedded in the new physical moisture sensor.

Based on the observations, the final conclusion of this part of the research is twofold:

- (i) the use of Functional Linear Models seems to be useful to identify the most useful predictors among the available signals;
- (ii) FLM could be used on data from different case studies with the purpose of verifying possible improvements in estimations and/or discover parts of signals which are useful and already unexplored for feature extraction with the MLR method. In the case of better estimation performance, FLM is a functional technique that can be taken into account also for the online implementation since the only information needed is how to combine the acquired signals with β -functions. In this case, clearly, the complexity of β -functions has to be consistent with the computational burden allowed on firmware.



(a) 4 acquisitions of signal 2 divided by nominal weight. Values in the first minutes of the cycle framed in red.



(b) Values in the first minutes of the cycle.

Figure 13.4: signal 2.

Part V

Conclusions

14

Limitations and future works

14.1 Limitations of the research

This chapter has the goal to be an objective assessment of the research activities conducted during my Ph.D. focusing on all the aspects related to the design of Soft Sensors in the industrial environment of fabric care R&D. This is my attempt to make the shortcomings as clear as possible together with the enhancement of the interesting results obtained during the process. Some next steps have already been expressed by people involved in this research and me; they have been summarized here as starting points for the near future.

When the main guidelines of fabric care modeling activities have been reported and discussed in this dissertation (1.5), 3 points have been included; now there's a need to linger on one among the others: the necessity of interpretable models. This is actually the common thread of all the methodologies and procedures exploited in this research. The proposal of interpretable models means to have interest and preference for methods for which the relation between inputs and outputs available is explained and the traceability is clear, e.g. in MLR method the role of each predictors is determined by the values of coefficients computed offline or in Symbolic Regression each selected model comes from combinations of predefined operands and operators representable with tree simple structures and it is possible to discover how each gene (or part of the final model)

contributes to the estimation. Basing on these considerations, and the fact that among the SSs considered, the main SS candidate for an online implementation is the one called poly3, the following shortcomings can be formalized:

1. the use of functional approaches described in part IV for the design of laundry weight SSs has not been extended also to the other case, i.e., SSs for laundry moisture estimation. This is identified as a methodological gap because functional methods and especially Function-on-Function regression can be applied to the moisture estimation problem with the same inputs and outputs used for the procedures discussed in this dissertation. The application of such techniques should be the missing piece of the scenario presented in this research and could represent the real benchmark concerning the performance achievable with interpretable methods. As expected, the constraint on the size of the dataset used for moisture estimation still applies to this case but an estimation error obtained with these procedures should be considered and evaluated.

In case of the availability of more data, functional regression could be useful also to perform offline simulation of drying cycles with the only aim of extending the already available tools used in fabric care R&D (introduced at the beginning of the dissertation in 1.1).

2. the most important shortcoming is the current lack of real simulations of the discussed methods online on machines. As reported in the description of the procedures, the use of Cross-Validations in the training phase of the proposed models applied to our real dataset was performed to limit the overfitting problem and hence taking into account the variability that comes from real situations. Clearly, without several real applications of the methods declared suitable for online implementation, it cannot be stated that solutions for the two main estimation problems have been reached. The results achieved have been obtained through offline simulations. Therefore, this research has actually provided answers about the feasibility of some kind of methodologies rather than others in the context of the development of data-driven SSs in resource-constrained industrial applications.

As a remark, the lack of results delivered by real online drying cycles can be explained with these two reasons

- (a) the need to test the candidates SSs with additional data produced by other different machines (WD for laundry moisture SS and TD for laundry weight SS); that is indeed the only way to test the existing machine-to-machine variability.

- (b) strict timings to require a first firmware version of the code for the SS candidate for the EoC goal (poly3).

14.2 Comparison with a priori hypotheses

In 4.4, some a priori hypotheses were advanced about the results and research answers of the entire project. Those hypotheses were based on the study made on the literature of SSs development and personal experience gained in several activities conducted in the fabric care R&D department during these last years. Now it's time to comment on those hypotheses in view of the results achieved. In the following, I'll review the list where each initial hypothesis reappears in *italics bold* together with the relative final retrospective consideration. This part is focused on *Why* question, i.e., the interest is on giving explanations for what has been achieved regardless of the fact that a given hypothesis has been confirmed or not.

- (a) ***the use of the mentioned methodologies allows model interpretability, i.e., it is supposed to be feasible to achieve a final precise comprehension of how available predictors or signals could be combined to achieve the goals. The signals that are used today by the current online implemented algorithms are supposed to be included in the set of statistically significant predictors selected by the discussed procedures.***

This supposition has been confirmed. In particular, MLR and SR methods helped to provide insights and suggestions on useful input information for the moisture estimation problem. They confirm the fact that linear combinations of some already available signals on WD-HP machines can be used to obtain results close to the target. The comprehension of how these predictors are combined is provided by the numerical values of the model coefficients which are available but omitted here. It has been confirmed that some statistically significant predictors for the moisture estimation problem are signals of the main temperatures and motor torque. Dealing with the other problem (laundry weight estimation), functional linear models can improve model interpretability and performance.

- (b) ***with the current information available online on machines it is possible to improve the current performances without additional physical sensors.***

This has not been satisfied completely. Dealing with error distributions, the median term can be lower than the one achieved with the current algorithm implemented online but the main problem is always the same: the variability of results that is

assumed to be due to the overfitting or scarcity of data. Some countermeasures have been adopted whenever possible, this kind of problem has been reduced but is still present in the poly3 model which is the SS candidate for the EoC goal using some controls on the estimated values. A fact that was not expected in the use of SR methodology for moisture estimation provided worse performance when compared to MLR. In simple terms, it is a little bit counter-intuitive that a smaller set of models performs better than a bigger one. One accredited hypothesis is the difficulty of governing the overfitting in the problem at hand, i.e., it is difficult to combine input signals in nonlinear combinations and at the same time explain the total amount of variability in the available data.

- (c) *the current feature engineering process involved in our already available SSs design process can be improved with the use of Functional Data Analysis techniques.*

This is correct. Functional Linear Regression allowed the comprehension of the role of a given signal used as a predictor for the estimation of the goal during time. This was observed studying the laundry load estimation problem and finding confirmations about the portions of given signals that are useful for predictions based on our experience.

- (d) *I am confident that also the calibration burden related to the current algorithm used for the automatic EoC in washer-dryer machines could be decreased using the data-driven SSs resulting from this research.*

This is supposed to be true but there is no evidence-based on real calibration procedure of the proposed algorithms. The procedure proposed as poly3 is so different from the current implemented solution discussed in 3.1. On one hand, the shortcoming of the current solution is the significant amount of time spent to set the thresholds; on the other hand, the proposed poly3 needs the training of several models offline: given a machine program (e.g. Cotton), it is necessary to train a moisture model for each load of interest (load A, B, C, etc). One important difference between the two procedures is that in poly3 the calibration phase can be performed automatically as long as the feature engineering process turns out to be robust; while the current approach is more subjective in the current way in which thresholds are set which depend on the operator/developer. For these reasons, experimental evidence is needed to answer this particular research question.

Summing up: Promising results have been achieved and a solution of the two main estimation problems discussed in this thesis seems to be achievable considering the

requirements and using data-driven SS. Nevertheless, considering the limitations of the research, only additional experimental evidence has the power to confirm the remaining hypotheses.

14.3 Future works

Some research directions have been defined as next steps:

- The work on the proposal of poly3 as a data-driven SS to address the EoC and TTE goals continues and evaluation online has been planned to fill the gap discussed above. An important aspect is that the activities planned for the laboratory and firmware developer to obtain evaluation on real tests depend on the status of several other projects, therefore, it is not possible to say for sure if this kind of evaluation will take place in the next 6 or 10 months because the projects focused on the current production have a higher priority.
- different kinds of techniques could also be taken into account to solve the problems at hand, such as the use of evolutionary fuzzy systems which constitute another possible methodology for the design of SSs considering the requirements we have. Some references under evaluation are for example: [Fazzolari, Alcalá, Nojima, Ishibuchi, and Herrera \(2012\)](#) and [Shapiro \(2005\)](#). Another method that is under evaluation to further lower the variability of the results in poly3 in the Bayesian Lasso technique ([Park and Casella \(2008\)](#)), a method that can be exploited in case of a limited availability of data to get more stable estimations. Just to mention other possibilities, a different sparse learner which is popular together with LASSO for SS design in process industry is the Relevance Vector Machine (RVM). This approach is compared with LASSO regularization in a recent work ([Urhan and Alakent \(2019\)](#)) and it could be used as an alternative sparse learner for the laundry moisture estimation problem.
- Functional Linear Models have been considered in general as a promising method to improve the current performance related to the laundry load estimation problem, its applicability could be useful for both washing and drying technology applications (in this thesis it has been analyzed only with data produced by dryer machines). Some other works in this direction are planned. In particular, the use of Functional Generalized Additive Models (FGAM) that is an approach I am considering taking inspiration from this work: [McLean et al. \(2014\)](#). FGAM allows more flexibility in the formulation of the functional problem because it

goes beyond the limiting assumption of linearity keeping the model interpretability that is a significant requirement for my research as stated. The interpretability is preserved by the additivity in the formulation and this is the most important characteristic of the methods seen in this dissertation to achieve the goals we have. The FGAM technique is under evaluation on the available data to obtain *beta*-surfaces which are very detailed representations of the role of predictors in estimations made possible by FGAM. An example has been provided in Figure 4.8 when we talked about the potential of FDA for our goals in this research. Figure 4.8 is only an example that comes from a preliminary use of FGAM; that's the reason why this method has not been treated in this dissertation.

Both of the problems considered in this dissertation are particularly interesting for applied research in fabric care R&D. As regards drying technology applications which are the object of my research, the present effort is on the proposal of different and efficient solutions for the EoC and TTE goals. The research explored in this thesis is considered as a starting point in this direction.

15

Conclusions

Continued advances in technology and data collecting nowadays are affecting also household major appliances. In this context, the main challenging goal is to produce more accurate, energy-efficient and user-friendly products in a resource-constrained environment like domestic appliances for fabric care.

This dissertation provides contributions to the field of drying technology for fabric care by introducing several methods for the determination of the main quantities of interest for optimal handling of the drying cycle: laundry weight and laundry moisture. In particular, as explained in part II, Soft Sensors are the proposed solutions chosen to avoid the use of costly physical devices or complex physical models. Several practical aspects of a real industrial environment are taken into account in this thesis: indeed, some issues related to data acquisitions, extraction, merging from different data sources, format, etc. have been addressed. Furthermore, a significant amount of time has been employed to acquire the different expertise necessary to fully understand tool routines and the timing of different related projects oriented to production. As regards the applied research topics, the contribution of this thesis is twofold:

1. proposal of a simple Soft Sensor for the online prediction of laundry moisture during drying cycles. This model allows the possibility to overcome some limitations due to the use of specific temperature sensors or calibration effort;

2. the proposal of functional methods to address the laundry load estimation task which is critical for the entire drying process in machines of our interest. This kind of approach allows the simplification of the current time-consuming feature engineering phase, offering an easier systematic way to fix the problem.

The core part of the dissertation is dedicated to the discussion and comparison of the proposed methods developed for the two tasks above. In particular, part III deals with SSs for laundry moisture content. Experimental data were used here to train models with available signals as predictors and laundry moisture content samples as the output in a supervised manner. As a final result, the best model, according to the requirements of this research, has been provided as a time series obtaining a *smooth* curve to predict the End-of-the-Cycle automatically. The proposal of this data-driven model, together with the related online implementation details, is actually the main research answer of this dissertation.

Similarly, part IV is entirely focused on the use of SSs for the estimation of dry laundry weight and in particular Functional Linear Models. The main benefit which comes from these methods for the purpose at hand in the short term is the improvement of the offline feature engineering process through the indication of the parts of the signals most influential for the output of interest; in the long term, these type of techniques could be useful online to exploit the information inside predictors and improve the estimation error itself. In general terms, data collecting for industrial applications can be expected to increasingly produce samples that are natural to analyze as functional data. Methods for functional data analysis will surely receive more and more attention in a variety of scientific fields as these types of data sets proliferate. Moreover, for the purpose at hand, the interpretability of the developed models is mandatory. Therefore, the use of simple (and especially additive) architectures was preferred when possible, reinforced by the fact that the number of tests in the available dataset is low. The additivity of used final models is the guideline of the methodologies and procedures experimented in this research; this common factor underlines how important this requirement is for the design of SSs in an industrial resource-constrained environment like the one described in this thesis.

Two strengths of this work are noteworthy: (i) the novelty: to the best of the author's knowledge there are no similar cases in literature concerning the online estimation of parameters of a moisture transfer model through porous media like laundry; (ii) an effective "soft" approach: no impact on hardware and no dedicated Design of Experiments (DOE) were required to conduct this industrially oriented research.

References

- Annual report. Technical report, JSPS PSE 143 committee, 2004. vol. 28.
- The simulation-driven design benchmark report. Technical report, Aberdeen Group, 2006.
URL <https://www.reden.nl/bestanden/AberdeenSimulationDrivenDesign.pdf>.
- New european energy label, 2012. URL http://www.newenergylabel.com/it/discover_the_label.
- Package refund, cran r documentation, 2018. URL <https://cran.r-project.org/web/packages/refund/refund.pdf>.
- Ahlberg J. H., Nilson E. N., and Walsh J. L.** The theory of splines and their applications. *Mathematics in Science and Engineering*, New York: Academic Press, 1967, 1967.
- Al-Obeidat F., Spencer B., and Alfandi O.** Consistently accurate forecasts of temperature within buildings from sensor data using ridge and lasso regression. *Future Generation Computer Systems*, 2018.
- Azeem M., Boughattas A., Wiener J., and Havelka A.** Mechanism of liquid water transport in fabrics; a review. *Fibres and Textiles (4) 2017 Vlákna a textil (4) 2017*, 58, 2017.
- Banks J.** *Handbook of simulation: principles, methodology, advances, applications, and practice*. John Wiley & Sons, 1998.
- Bastin G.** *On-line estimation and adaptive control of bioreactors*, volume 1. Elsevier, 2013.
- Becker M. C., Salvatore P., and Zirpoli F.** The impact of virtual simulation tools on problem-solving and new product development organization. *Research Policy*, 34 (9):1305–1321, 2005.
- Berger J., Gasparin S., Dutykh D., and Mendes N.** Accurate numerical simulation of moisture front in porous material. *Building and Environment*, 118:211–224, 2017.
- Berk R. A.** *Statistical learning from a regression perspective*, volume 14. Springer, 2008.
- Berk R. A. and Freedman D. A.** Statistical assumptions as empirical commitments. *Law, punishment, and social control: Essays in honor of Sheldon Messinger*, 2:235–254, 2003.
- Bishop C. M.** *Neural networks for pattern recognition*. Oxford university press, 1995.

- Blauert P., Boldt F., Wobkemeier M., Pilgram P., and Rehfuess U.** Method for determining the mass of wet laundry in a laundry drum, April 16 1996. URL <http://www.google.mw/patents/US5507054>. US Patent 5,507,054.
- Bollinger G.** Book review: Regression diagnostics: Identifying influential data and sources of collinearity, 1981.
- Chbat N. W. and Badami. V. V.** Systems and methods for determining drying time for a clothes dryer, 2000.
- Cook R. D. and Weisberg S.** *Applied regression including computing and graphics*, volume 488. John Wiley & Sons, 2009.
- Crank J.** *The mathematics of diffusion*. Oxford University Press, 1979.
- Didier L. and Geay J.** Washing and/or drying machine having means for determining a load of clothes in a rotary drum, May 3 1988. URL <https://www.google.co.in/patents/US4741182>. US Patent 4,741,182.
- Dinçer İ. and Zamfirescu C.** *Drying phenomena: theory and applications*. John Wiley & Sons, 2016.
- Dincer I. and Zamfirescu C.** *Drying phenomena: theory and applications*. John Wiley & Sons, 2016.
- Domingos P.** The role of occam’s razor in knowledge discovery. *Data mining and knowledge discovery*, 3(4):409–425, 1999.
- Doraiswami R. and Cheded L.** Robust model-based soft sensor: design and application. *IFAC Proceedings Volumes*, 47(3):5491–5496, 2014.
- Fazzolari M., Alcala R., Nojima Y., Ishibuchi H., and Herrera F.** A review of the application of multiobjective evolutionary fuzzy systems: Current status and further directions. *IEEE Transactions on Fuzzy systems*, 21(1):45–65, 2012.
- Fernandez-Delgado M., Sirsat M., Cernadas E., Alawadi S., Barro S., and Febrero-Bande M.** An extensive experimental survey of regression methods. *Neural Networks*, 2018.
- Ferraty F.** *Recent advances in functional data analysis and related topics*. Springer Science & Business Media, 2011.

- Ferraty F.** Regression on functional data: methodological approach with application to near-infrared spectrometry. 2014.
- Ferraty F., Laksaci A., and Vieu P.** Estimating some characteristics of the conditional distribution in nonparametric functional models. *Statistical Inference for Stochastic Processes*, 9(1):47–76, 2006.
- Fisher R. A.** Statistical methods for research workers. In *Breakthroughs in statistics*, pages 66–70. Springer, 1992.
- Fox J.** *Regression diagnostics: An introduction*, volume 79. Sage, 1991.
- Funke C. and Jonsson D.** A framework for implementing simulation-driven design, 2019.
- Gasparin S., Berger J., Dutykh D., and Mendes N.** Advanced reduced-order models for moisture diffusion in porous media. *Transport in Porous Media*, 124(3): 965–994, 2018.
- Goldsmith J., Bobb J., Crainiceanu C. M., Caffo B., and Reich D.** Penalized functional regression. *Journal of Computational and Graphical Statistics*, 20(4):830–851, 2011.
- Goodfellow I., Bengio Y., and Courville A.** *Deep learning*. MIT press, 2016.
- Goodwin G. C.** Predicting the performance of soft sensors as a route to low cost automation. *Annual Reviews in Control*, 24:55–66, 2000.
- Haghi A. and Ghanadzadeh H.** A study of thermal drying process. *Indian Journal of Chemical Technology*, 12(6):654–663, 2005.
- Halevy A., Norvig P., and Pereira F.** The unreasonable effectiveness of data. 2009.
- Hamdaoui M., Baffoun A., Chaaben K. B., and Hamdaoui F.** Experimental study and mathematical model to follow the drying phenomenon of knitted textile fabric. *Journal of Engineered Fibers and Fabrics*, 8(3):155892501300800308, 2013.
- Harrell Jr F. E.** *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer, 2015.
- Hastie T., Tibshirani R., Friedman J., and Franklin J.** The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, 27(2): 83–85, 2005.

- Ho Y.-C. and Pepyne D. L.** Simple explanation of the no-free-lunch theorem and its implications. *Journal of optimization theory and applications*, 115(3):549–570, 2002.
- Hoaglin D. C. and Welsch R. E.** The hat matrix in regression and anova. *The American Statistician*, 32(1):17–22, 1978.
- Horváth L. and Kokoszka P.** *Inference for functional data with applications*, volume 200. Springer Science & Business Media, 2012.
- Jang J., Sun C., and Mizutani E.** Neuro-fuzzy and soft computing upper saddle river, 1997.
- Jolliffe I. T.** Springer series in statistics. *Principal component analysis*, 29, 2002.
- Jordaan E., Kordon A., Chiang L., and Smits G.** Robust inferential sensors based on ensemble of predictors generated by genetic programming. In *International Conference on Parallel Problem Solving from Nature*, pages 522–531. Springer, 2004.
- Kadlec P., Gabrys B., and Strandt S.** Data-driven soft sensors in the process industry. *Computers & Chemical Engineering*, 33(4):795–814, 2009.
- Kalivas J. H.** Two data sets of near infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, 37(2):255–259, 1997.
- Kano M. and Fujiwara K.** Virtual sensing technology in process industries: trends and challenges revealed by recent industrial applications. *Journal of chemical engineering of Japan*, page 12we167, 2012.
- Kronberger G.** *Symbolic Regression for Knowledge Discovery - Bloat, Overfitting, and Variable Interaction Networks*. PhD thesis, Technisch-Naturwissenschaftliche Fakultät, 2010.
- Kronberger G.** Symbolic regression for knowledge discovery: bloat, overfitting, and variable interaction networks. *ACM SIGEVOlution*, 5(4):25–25, 2011.
- Kucuk H., Midilli A., Kilic A., and Dincer I.** A review on thin-layer drying-curve equations. *Drying Technology*, 32(7):757–773, 2014.
- Lotov A. V. and Miettinen K.** Visualizing the pareto frontier. In *Multiobjective optimization*, pages 213–243. Springer, 2008.
- Luikov A. v.** *Analytical heat diffusion theory*. Elsevier, 2012.

- McLean M. W., Hooker G., Staicu A.-M., Scheipl F., and Ruppert D.** Functional generalized additive models. *Journal of Computational and Graphical Statistics*, 23(1):249–269, 2014.
- Miller T.** Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 2018.
- Mitchell T.** Machine learning, mcgraw-hill higher education. *New York*, 1997.
- Molnar C.** *Interpretable Machine Learning*. 2019. <https://christophm.github.io/interpretable-ml-book/>.
- Mourtzis D., Doukas M., and Bernidaki D.** Simulation in manufacturing: Review and challenges. *Procedia CIRP*, 25:213–229, 2014.
- Müller H.-G.** Statprob: The encyclopedia sponsored by statistics and probability societies, chapter functional data analysis, 2011.
- Negahban A. and Smith J. S.** Simulation for manufacturing system design and operation: Literature review and analysis. *Journal of Manufacturing Systems*, 33(2): 241–261, 2014.
- Noguerales J. L. T.** *On the theory and practice of variable selection for functional data*. PhD thesis, PhD thesis, Universidad Autónoma de Madrid, 2015.
- Oehlert G. W.** A few words about reml. *Verfügbar unter: http://users.stat.umn.edu/~gary/classes/5303/handouts/REML.pdf*. Abrufdatum, 29:2014, 2014.
- Ogden R. T. and Goldsmith J.** Functional data analysis: Techniques and applications. Technical report, Columbia University, 2014. URL https://www.mailman.columbia.edu/sites/default/files/pdf/fda_shortcourse.pdf.
- Ookita K.** Operation and quality control for chemical plants by soft-sensors. *CICSJ Bulletin*, 24:31–33, 2006.
- Orzechowski P., La Cava W., and Moore J. H.** Where are we now?: a large benchmark study of recent symbolic regression methods. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 1183–1190. ACM, 2018. Kyoto, Japan.
- Osborne B.** Applications of near infrared spectroscopy in food analysis. *Near infrared spectroscopy in food analysis*, pages 117–161, 1986.

- Pan N. and Gibson P.** *Thermal and moisture transport in fibrous materials*. Woodhead Publishing, 2006.
- Park T. and Casella G.** The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686, 2008.
- Pelczer A.** Method for determining the load in a washing machine, November 20 2012. US Patent 8,312,581.
- Prasad V., Schley M., Russo L. P., and Bequette B. W.** Product property and production rate control of styrene polymerization. *Journal of Process Control*, 12(3): 353–372, 2002.
- Ramsay J. O. and Dalzell C.** Some tools for functional data analysis. *Journal of the Royal Statistical Society: Series B (Methodological)*, 53(3):539–561, 1991.
- Ramsay J. O. and Silverman B. W.** *Applied functional data analysis: methods and case studies*. Springer, 2007.
- Ramsay J., Hooker G., and Graves S.** Introduction to functional data analysis. In *Functional data analysis with R and MATLAB*, pages 1–19. Springer, 2009.
- Ramsay J. and Silverman B.** Principal components analysis for functional data. *Functional data analysis*, pages 147–172, 2005.
- Rish I. and Grabarnik G.** *Sparse modeling: theory, algorithms, and applications*. CRC press, 2014.
- Ruppert D.** Selecting the number of knots for penalized splines. *Journal of computational and graphical statistics*, 11(4):735–757, 2002.
- Sander A. and Kardum J. P.** Experimental validation of thin-layer drying models. *Chemical Engineering & Technology: Industrial Chemistry-Plant Equipment-Process Engineering-Biotechnology*, 32(4):590–599, 2009.
- Scholkopf B. and Smola A. J.** *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2001.
- Searson D. P.** Gptips 2: an open-source software platform for symbolic data mining. In *Handbook of genetic programming applications*, pages 551–573. Springer, 2015.
- Shapiro A. F.** Fuzzy regression models. *Article of Penn State University*, 102(2): 373–383, 2005.

- Sharma S. and Tambe S. S.** Soft-sensor development for biochemical systems using genetic programming. *Biochemical Engineering Journal*, 85:89–100, 2014.
- Shephard M. S., Beall M. W., O’bara R. M., and Webster B. E.** Toward simulation-based design. *Finite Elements in Analysis and Design*, 40(12):1575–1598, 2004.
- Smargiassi E., Potena G. P., Minni F., Concettoni F., and Animobono V.** Method for estimating the weight of the contents of a household appliance, April 30 2013. US Patent 8,432,160.
- Soares S. G.** *Ensemble learning methodologies for soft sensor development in industrial processes*. PhD thesis, Departamento de Engenharia Electrotécnica e de Computadores da Faculdade de Ciências e Tecnologia da Universidade de Coimbra, 2015.
- Solomatine D. P. and Ostfeld A.** Data-driven modelling: some past experiences and new approaches. *Journal of hydroinformatics*, 10(1):3–22, 2008.
- Thomke S. H.** Simulation, learning and r&d performance: Evidence from automotive development. *Research Policy*, 27(1):55–74, 1998.
- Ullah S. and Finch C. F.** Applications of functional data analysis: A systematic review. *BMC medical research methodology*, 13(1):43, 2013.
- Urhan A. and Alakent B.** An exploratory analysis of biased learners in soft-sensing frames. *arXiv preprint arXiv:1904.10753*, 2019.
- van Oudenaarden A.** The importance of diffusion and gradients for cellular regulation; course of systems biology. Technical report, MIT, 2009. URL http://web.mit.edu/biophysics/sbio/PDFs/L15_notes.pdf.
- Vanneschi L., Castelli M., and Silva S.** Measuring bloat, overfitting and functional complexity in genetic programming. In *Proceedings of the 12th annual conference on Genetic and evolutionary computation*, pages 877–884. ACM, 2010.
- Vanneschi L. and Poli R.** Genetic programming—introduction, applications, theory and open issues. *Handbook of natural computing*, pages 709–739, 2012.
- Vapnik V. and Vapnik V.** Statistical learning theory wiley. *New York*, pages 156–160, 1998.
- Welch G. and Bishop G.** An introduction to the kalman filter. university of north carolina. In *Department of Computer Science, TR*. 1995.

- Winkler S. M.** *Evolutionary system identification: modern concepts and practical applications*. Trauner, 2008.
- Winkler S. M., Affenzeller M., and Wagner S.** Using enhanced genetic programming techniques for evolving classifiers in the context of medical diagnosis. *Genetic Programming and Evolvable Machines*, 10(2):111–140, 2009.
- Xu Q.-S. and Liang Y.-Z.** Monte carlo cross validation. *Chemometrics and Intelligent Laboratory Systems*, 56(1):1–11, 2001.
- Yan X. and Su X.** *Linear regression analysis: theory and computing*. World Scientific, 2009.

Acknowledgments

Writing acknowledgments is never easy since one undergoes the risk of forgetting somebody, but, first of all, I want to thank my advisors Gian Antonio Susto, Alessandro Beghi and Mirco Rampazzo who have shown extraordinary patience and generosity, without mentioning their professional experience. I am very grateful to Gian Antonio and Mirco for everything they have done for me during these years.

I would like to deeply thank my supervisors at PUC-PR in Curitiba (Brazil) for their availability, guide, patience, and friendliness; thank you Prof. Gilberto Reynoso- Meza and Prof. Leandro dos Santos Coelho.

I want to thank the Electrolux GTD CA&CCF team for their support and help; it has been a pleasure to be part of you during these years. Special thanks to Terenzio Girotto for his patience and professional guide; his help and kindness has always been extremely useful for me.

One thank goes to Prof. Gilberto Reynoso-Meza, Seán McLoone and all the committee that patiently read and reviewed this thesis.

The most special thank goes to my mother, Maria Silva, she always supported me giving me an example of life and she taught me the love for what in life really matters; this thesis is dedicated to her.

The last, special and unique thank goes to Bianca Andrade Colle, the person who has supported me every day during this last year. She had the strength to patiently wait for me and her empathy has been crucial in many situations. She is the meaning in my life and she helps me to grow up and eventually be proud of what I am now.

I remember also my colleagues in the company and university as well, for the pleasant (always interesting) conversations of these last years.

Thanks to my friends Stefano Zorzi, Precious Ugo Abara and Marco Rossi. Life without them would be boring and partly meaningless.

To conclude, I want to express my recognition to who has believed in me and especially to who has not. The former for their regard and commitment, the latter because they forced me to improve myself.