Data-Driven Models for the Determination of Laundry Moisture Content in a Household Laundry Treatment Dryer Appliance

Giuliano Zambonin[1,2], Fabio Altinier[2], Alessandro Beghi[1], Leandro dos Santos Coelho[3], Terenzio Girotto[2], Mirco Rampazzo[1], Gilberto Reynoso-Meza[3], and Gian Antonio Susto[1]

[1] Department of Information Engineering, University of Padova, Italy
zambonin@dei.unipd.it
[2] Electrolux Italia S.p.a., Porcia (PN), Italy
[3] Industrial and Systems Engineering Graduate Program (PPGEPS), Pontificia Universidade Católica do Paraná (PUC-PR), Curitiba, Brazil.

Abstract. Two methods based on Regression are presented to determine the moisture content of items, e.g. clothes and the like, which are introduced in a household laundry dryer appliance. The aim of this work is to develop Soft Sensors (SS) for a household Heat Pump Washer-Dryer (WD-HP) to provide an estimation of the desired signal (the laundry moisture during drying) avoiding the use of additional physical sensors with the goal of improving the current performance in terms of precision and energy consumption of the automatic drying cycle and using the machine equipment already available.
On an algorithmic point of view, the SS developed in this work exploits regularization methods and Genetic Programming for Symbolic Regression in order to find suitable models for the purpose at hand. Proposed approaches have been tested on real data provided by an industrial partner.

Keywords: Domestic Appliances, Fabric Care, Genetic Programming, Heat Pump Washer-Dryer, Machine Learning, Moisture Transfer Models, Soft Sensors, Symbolic Regression.

1 Introduction

Household appliance manufactures are nowadays competing to provide more accurate, resources-efficient and user-friendly products. One of the main obstacles in optimizing household appliances processes can be related to the uncertainty in the laundry loaded in the appliance; the laundry characteristics (like weight, fabric, contained water) have a major impact in the drying and washing processes: being aware or estimating such characteristics can enable several process optimizations, in terms of both performance and consumption measurements.

The usage of dedicated physical sensors to characterize laundry is generally not possible or not costly-effective in household appliances. For these reasons, manufacturers have to resort to indirect information on the laundry inferred from other sensors or provided by the users. In this perspective, Soft Sensing [1] technologies may provide a viable solution. SS are statistical model-based technologies used in industrial environments to provide an estimate of quantities that may be unmeasurable or costly/time-consuming to measure, based on more accessible variables. SS exploits already in-place sensors/information, therefore representing a cost-free solutions for improving product/process performances. Such technology is generally based on *Machine Learning* (ML) supervised techniques [2] 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 process industry can be found in [3] 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. Another essential reference can be found in [4] which presents some ensemble learning methodologies for SS development in industrial processes.

Few ML-based solutions for fabric care appliances are available in the literature; for example in [5, 6, 7, 8]. SS based on Machine Learning approaches have been presented to estimate the laundry weight in washing machines and we believe that this may be traced back to two main issues in developing SS for fabric care appliances: (i) the effort in collecting a sufficient amount of laboratory data, where the load weight and water content is accurately measured; (ii) the complexity of embedding a ML-based solutions in household equipment. In this work we overcome these two issues by exploiting laboratory data already collected for other product development purposes and by exploiting *regularization* [9], a Machine Learning framework that allows, in some cases, to provide effective sparse linear models that are easily implementable.

The mainly original contribution of this paper is the proposal of a model starting from real data for the online estimation of laundry moisture during drying cycles in household dryer appliances (the case study is a Washer-Dryer machine). An in-dept study has been done considering available works in literature for different drying applications; moreover, Symbolic Regression has been used to determine appropriate models for the goal with the assumption to use entire available signals as predictors.

The rest of this paper is organized as follows. Section 2 is dedicated to introduce the problem and to review SS solutions and estimation approaches in fabric care appliances; in Section 3, two methods for the estimation of the laundry moisture content during drying are illustrated. In Section 4 the description of the available dataset and experimental settings for each of the method tried is given. Section 5 is devoted to summarize the experimental results obtained on real industrial data. Concluding remarks are then reported in Section 6.

2 Related works

In simple terms, we could describe the physical process at hand (laundry drying cycles) as a phenomenon in which the transfer of moisture from a porous media (laundry) to the environment (the drum) takes place, in particular drying is a process of simultaneous heat and moisture transfer which induces changes in the product undergoing dehydration.
Heat and mass transfer in porous media is a complicated phenomenon; Scheidegger [10] claimed that the structure of porous media is too complex to be described precisely either in micro-scale or macro-scale. The mechanism of heat and mass transfer in moist porous media is so sophisticated that Luikov [11] attributed all the transport phenomena to the effects of temperature, moisture content, and pressure from the macro-phenomenon viewpoint by applying irreversible thermodynamics.
In the literature, the study of textile drying is limited, however, there are numerous studies of the technique of drying of foods, (for example [12, 13, 14]). In particular kucuk et al. [13] proposes a comprehensive review of drying curve models available and their comparisons for several applications while Younis et al. [14] is a recent work which makes comparisons between thirteen different mathematical models with non-linear regression analysis for describing the garlic drying process. Talking about food preservation, the mechanism of moisture transfer in food is complex and very often diffusion models (e.g., [15]) are given significant attention in the literature due to ease of formulation.
One of the challenges for drying research still is the incorporation of the knowledge of basic thermodynamics and transport phenomena into the description of phase equilibria and drying kinetics [16]. Although many theoretical and experimental drying studies have been undertaken by many researchers (e.g.,[17]) to predict/determine moisture transfer, particularly drying profiles of various products, some models on moisture transfer parameters are available in the literature (e.g., [18]), with a wide variation of reported values, due to the complexity of the products and methods of estimation [19].
The goal of the present work is to select a suitable model to describe the laundry moisture evolution during drying process of our interest (in a household Washer-Dryer machine) with the aim of improving performances being more accurate in the automatic determination of the End-of-the-Cycle (EoC) which correspond to the instant in which the laundry is perfectly dryed[4].
Considering the difficulties described above and the variability due to several possibilities for the laundry composition inside the drum during drying cycles in machines of our interest (WD or household dryers as well), the challenge is not

[4]An important expected improvement is also the possibility to get better estimations of how much time is needed to stop the cycle at a predefined moisture target (Time-to-End) from the instant in which the model is available, e.g. 20 minutes from the beginning. Such an estimate can be used to provide an information to the user (through the user interface) which is more accurate than the one provided now on machines.

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 and data-driven techniques.
Significant references for the problem at hand in this work are [20, 21]. Haghi et al. [20] provide an experimental examination of the convective drying behaviour of wool; using the acquired data, the aim is to assess the ability of some selected drying models from literature to quantify the moisture removal behaviour, especially for wool convective drying. In [21] the applicability of three mathematical models for the description of the drying kinetics was investigated for describing the drying kinetics of materials of various origins and of the inner structure dried in four different dryers. One of the most significant references for the work presented here is [22] 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 our industrial partner; these results were found independently.

3 Modelling

In this section a description of methods used to obtain models for moisture transfer during drying is presented; in particular moisture transfer through laundry during drying is the process of interest and the first part is a basic explanation of this phenomenon reporting the main references from literature.

3.1 Moisture transfer modelling by diffusion

Here an introduction to the physical description of the process of interest will be given without aspiration of fully explain the subject (interested reader will find excellent references provided below). A drying process always involves moisture transfer which is essentially driven by heat and mass transfer; heat and mass transfer are governed by similar equations of diffusion (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 [23].

Moisture diffusion is the process during which water molecules migrate through given materials. When we are only interested in mono-component mass transfer, i.e. water, the diffusion process is quite similar to the thermal conduction process [23]. Moisture transfer through porous media (like laundry) can be therefore simply explained according to moisture concentration gradient variation using the same equations exploited for heat and mass transfer ([24] and [25]). Then, the most simple equation that can be used to describe the process of interest is the one known as Fick's general second law of diffusion:

$$\frac{\partial \phi(x,t)}{\partial t} = D \frac{\partial^2 \phi(x,t)}{\partial x^2} \tag{1}$$

where t is the time in $[s]$, x is the spatial coordinate ($[m]$) (gradient direction through the porous medium), D is the diffusion coefficient $[m^2/s]$ and ϕ is the particle concentration of water $[kg/m^3]$. From these types of mathematical solution, the following general model for transient moisture transfer was proposed as a simplified solution by Dincer & Dost ([23]):

$$\phi = LFe^{-St} \tag{2}$$

where LF is called Lag Factor (adimensional) and S is the moisture transfer coefficient $[s^{-1}]$; drying coefficient shows the drying capability of an object or product per unit time and Lag Factor is an indication of internal resistance of an object to the heat and/or moisture transfer during drying. This simplified solution is widely used in literature to describe drying processes. The model provided by equation (2) is the most simple 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. For this reason it is used here as a reference and starting point for further analysis based on real data available from machines in order 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 our study on physical models for diffusion processes is reported here: [26], it contains details about the solution of the aforementioned equation (1) as well.

Here it is pointed out that, for the rest of the paper our moisture reference is computed in percentage according to the following relation: $y[\%] = \frac{(weight - w_c)}{w_c} 100[\%]$ where y is used for moisture, $weight$ is the laundry weight (in Kg) signal obtained from a balance positioned under the machine during drying cycles, after filtering and rescaling[5] procedures while w_c is the conditioned weight, i.e. the weight of dry laundry. Besides, the desired target for improvements of automatic EoC is to keep the moisture error (RMSE explained below) under the threshold of 3% for available data (Cotton laundry).

3.2 Data-Driven methods

Method 1: Genetic Programming for Symbolic Regression This kind of approach has been used with the aim of 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 kind of operations specified by the user could be used to process input signals

[5]Details about computed features, will be omitted here because of intellectual property rights.

to obtain the output reference (which is a reconstruction from the signal that comes from the balance under the machine during drying tests in laboratory). The focus has been the development of this method without a priori assumptions on the model (linear or polynomial form etc.) in order to discover suitable models for our purpose although hardly representable on a firmware version.

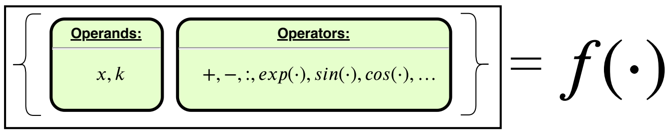


Fig. 1: Representation of GP-based symbolic regression; $\mathbf{y}$=laundry moisture content during drying cycle, $\mathbf{x}$=independent variables (signals available), k= parameters

Genetic Programming (GP) is a supervised learning method of the evolutionary computation field motivated by an analogy to biological evolution. GP creates successor hypotheses by repeatedly mutating and cross-overing parts of the current best hypotheses, with expectation to find a good solution in the evolution process. A symbolic regression problem consists in finding the symbolic function that matches a given set of data as closely as possible (Fig. 1). By training the Genetic Programming algorithm with the given data set, the relationship between the input and output is represented by functions generated in the training process. If the error rate reached a certain threshold, the training can be stopped and the testing can be applied to verify the effectiveness of the best function. This method is usually used to discover governing equations from noisy measurement data [27]; 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 here: [28].

In GP approach a population of computer programs is developed. The primary genetic operations that are used to create new programs from existing ones are:

- **Crossover:** The creation of a child program by combining randomly chosen parts from two selected parent programs.
- **Mutation:** The creation of a new child program by randomly altering a randomly chosen part of a selected parent program.

GP departs significantly from other evolutionary algorithms in the implementation of the operators of crossover and mutation.

Symbolic Regression has been tested using a GP tool called GPTIPS (Genetic Programming Toolbox for the Identification of Physical Systems, [29]) for use with MATLAB®; it employs a unique type of symbolic regression called multigene Symbolic Regression that evolves linear combinations of non-linear transformations of the input variables.

In the first generation of the algorithm, a population of random individuals is generated. For each new individual, a tree representing each gene is randomly generated (subject to depth constraints) using the users specified palette of building block functions and the available M input variables $x_1, \ldots, x_M$. In the first generation the algorithm attempts to maximize diversity by ensuring that no individuals contain duplicate genes. However, due to computational expense, this is not enforced for subsequent generations of evolved individuals. Each individual is specified to contain (randomly) between 1 and G_{max} genes (trees). G_{max} is a parameter set by the user.

Method 2: Polynomial model for drying cycle prediction The approach discussed here deals with a mathematical model developed to provide a smooth description of the drying phenomenon of fabrics basing on our available data. Differently from previous method the structure of the model is fixed 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).

Fig. 2 shows an explanation of the proposed procedure in details. The available dataset has been used offline to select the best model for our purpose and it was done checking performance of fitting 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 to consider some well-known indices used for goodness-of-fit analysis: Root Mean Square Error (RMSE) and coefficient of determination[6] (R^2), 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 was found in a 3^{rd} degree polynomial model in time (t) and parameters a, b, c, d:

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

this is the structure that provides the best fit performance on our data (WD HP case study) and is proposed here as the model to explain the evolution of our drying cycles of interest. After our 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

[6]RMSE is represented in the next Section; the coefficient of determination is the square of the correlation coefficient; a higher coefficient of determination is an indicator of a better goodness of fit for the observations.

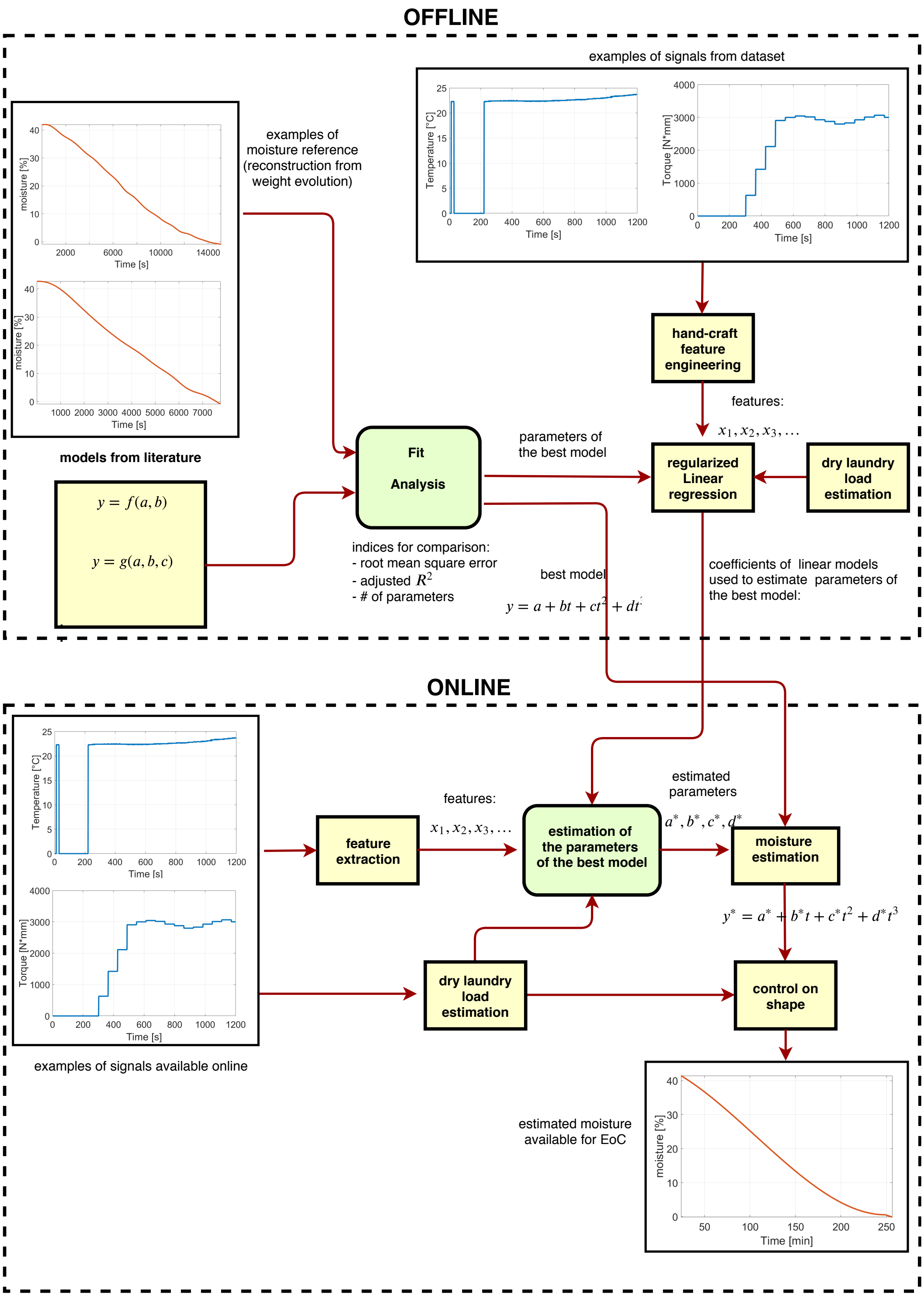


Fig. 2: Visual explanation of the algorithm proposed as Method 2

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 the expert knowledge and are fixed. The hand-craft features[7] used 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.
An essential information that is exploited here is the one related to the (dry) weight of the laundry [dry laundry load estimation block in Fig. 2 OFFLINE part], which is a value that is estimated online using the available information of the inertia of the laundry weight. So we use the information of the laundry weight quantity 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.

Table 1: Drying models fitted to experimental data; t =time, y =moisture

Model	**expression**	**RMSE[%]**	$\mathbf{R^2}$	**#params**
Lewis	$y = e^{-at}$	9.069	0.781	2
Henderson & Pabis	$y = ae^{-bt}$	1.147	0.996	3
quadratic	$y = a + bt + ct^2$	1.051	0.997	3
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 2 summarises the results of fitting between moisture references and several models used for different drying applications (see also [13, 20]); this study is represented as "Fit Analysis" block in Fig. 2 ONLINE 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; indeed, 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 in this step. The online part is executed on the machine (WD-HP) and it deals

[7]Details will be omitted here because of intellectual property rights.

with some simple passages: the first is the feature extraction process which is performed computing the features defined in the offline part after the first part of the cycle (20 minutes from the start) in order to have time to collect useful information from signals; the second passage 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 Fig. 2 ONLINE part]: once the estimation of the (dry) laundry load weight is available[8], 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:

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

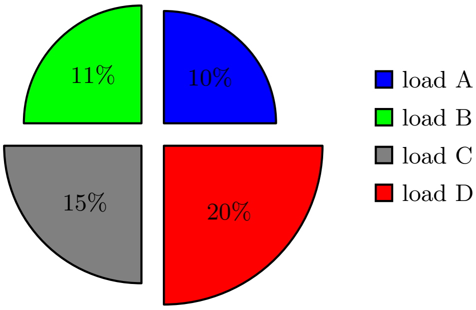
where $h_i, i = 1, \ldots, 7$ are the features selected offline and $\alpha_i, \beta_i, \gamma_i, \delta_i; i = 0, \ldots, 7$ are linear regression coefficients.
We point out that the computation of the drying rate which is the derivative of the moisture content in time, can be useful to detect anomalous estimation of laundry and this fact is exploited in Method 2 to modify the estimation in case; this situation is handled as a 'control on shape'[9] of the polynomial model (Fig. 2).

4 Experimental Setting

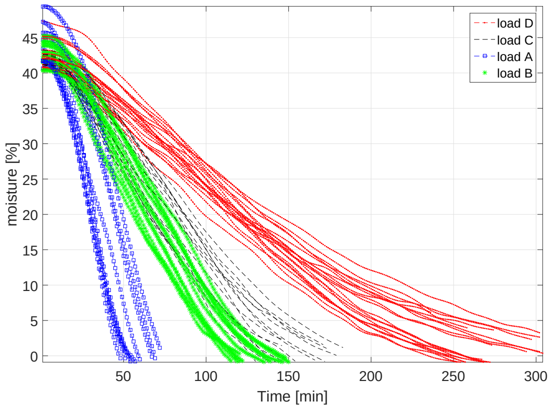
Details about the available dataset are reported here. In particular Fig. 3a and Fig. 3b show distribution of the available drying tests in terms of nominal weight (weight of dry laundry); in particular in Fig. 3b the output reference is depicted (i.e. the laundry moisture evolution during drying obtained from the evolution of weight of the wet laundry during drying). Time series represented here were used as output reference to train models. It's easy to see that moisture references are very different comparing small and large loads. A number of 56 drying tests are available for our development and 29 signals acquired for each observation.

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

[9] This simple control consists on a simple fixed decrease of the moisture 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).



(a)



(b)

Fig. 3: Data visualization by (dry) laundry load weight: **(a)** Laundry load weight distribution in the data, 4 classes in [kg]; **(b)** laundry moisture reference

As regards the methods introduced before, talking about Symbolic Regression, some choices have been made in order to ensure the convergence of the optimization procedure implemented in the used tool and to reduce the overfitting risk due to models that are not suitable for test data.
In particular our decisions for configuration setting is summarized and explained below: (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 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 together in the multigene approach; in this case the value has been fixed to 5 (low value) 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. The stopping criterion is related to the number of generations (ii); we used a high value here (1000) in order to be more confident about the convergence of the algorithm.
An entire moisture time series is selected randomly at each iteration to belong to a 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: $\sqrt{\frac{1}{n_i}\sum_{t=1}^{n_i}(y_i(t) - \hat{y}_i(t))}$ 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 Section 5). Estimation error is always computed considering th last part of the cycle: under 20% of reconstructed moisture because it is the part of interest in the application for the EoC.

In Method 2 linear models used to obtain the parameters a,b,c and d for the model summarized in equation (3) were trained using regularized linear regression with LASSO [2] which is exploited for its sparsity to select a set of statistically meaningful predictors (the features mentioned in Section 5). In this phase, the 70% of available observations for each load are used as the training set, while, the 30% are used for test in 100 Monte Carlo Cross Validations. All results also in this case appear in terms of error distributions and visualized using boxplots (Section 5).

As regards inputs and outputs of the 2 proposed methods, the output is the laundry moisture content value for both, while inputs are signals already available on machine and the difference between the procedures consists in the processing of signals: in particular, Method 1 exploits signals as they are online and combine them offline in linear and nonlinear ways in order to determine a suitable model (the model with the lowest RMSE); once the model has been fixed, it's necessary to evaluate its complexity for an implementation online. On the other hand, in Method 2 another input is the time elapsed from the beginning of the cycle (in $[s]$); some predefined features are computed online from available signals and used as predictors in linear regression models which compute the parameters of a 3^{rd} degree polynomial in time that provides the laundry moisture predicted value.

5 Results

Method 1: Genetic Programming for Symbolic Regression Fig. 4 refers to Method 1 (Symbolic Regression) and shows results obtained so far as moisture error in percentage 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.

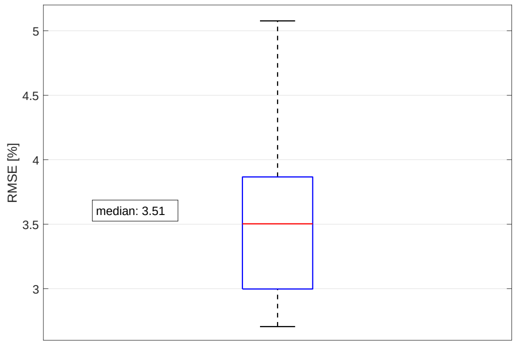


Fig. 4: Current performance in terms of RMSE distribution (in [%] of moisture content) using Symbolic Regression with GP approach using all input signals available

The obtained solutions are visualized in distribution in Fig. 4, but each iteration could give a different result in terms of the structure of the selected model, e.g. one solution can be the following:

$$\begin{aligned}\hat{\mathbf{y}} =& k_0 + k_1 \mathbf{x_2}^3 + k_2 \mathbf{x_7} \mathbf{x_2} + k_3 \mathbf{x_3} + k_4 \mathbf{x_7} + k_5 \mathbf{x_8} \\ &+ k_6 \mathbf{x_{16}} + k_7 \mathbf{x_{21}} + k_8 \mathbf{x_{23}} + k_9 \mathbf{x_{26}} + k_{10} \mathbf{x_{27}}\end{aligned}$$

where $k_i, i = 1, \ldots 10$ are found coefficients[10] and $\mathbf{x_i}$ are signals.

The solution selected at the next iteration could show different powers of involved input signals, but the goal here is the discovery of the most useful models for our purpose i.e. which structures have been selected more frequently in all iterations (if any). It turns out that the most selected structures are the ones which involve main signals of temperatures and motor torque (or signals that are elaboration of those); this suggests that these are also the meaningful signals for our goal.

[10]Details about computed coefficients and solution, will be omitted here because of intellectual property rights.

Method 2: Polynomial model for drying cycle prediction In Fig. 5 results in terms of error distribution are represented. It shows that in this case the variability of the error is high and its due to the variability of the estimation of parameters of the polynomial model, nevertheless, the focus of our work now is on the reduction of this variability. In particular Fig. 5 shows results for each laundry

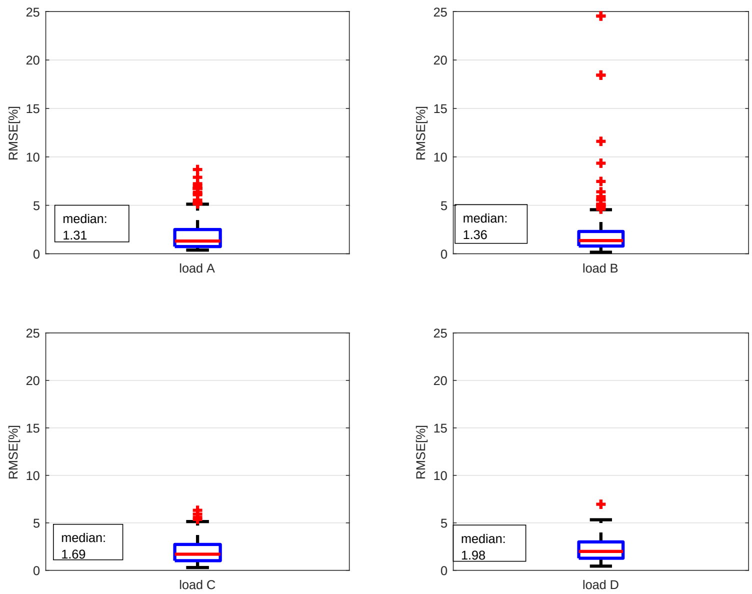


Fig. 5: Current performance in terms of RMSE distribution (in [%] of moisture content) using Method 2. Results divided by laundry nominal weight

load amount considering the fact that the algorithm exploits a load estimation phase (as represented in Fig. 2). As depicted in Fig. 5 median term in 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 EoC in Cotton laundry typology which is the case study at hand, but the current problem which is still present is the error variability as explained before. Fig. 6 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 preferable as a solution because of its smoothness; in fact in this case inputs are not rough signals from the acquisitions and once the estimation of parameters of the model is ready the result is a 3^{rd} degree

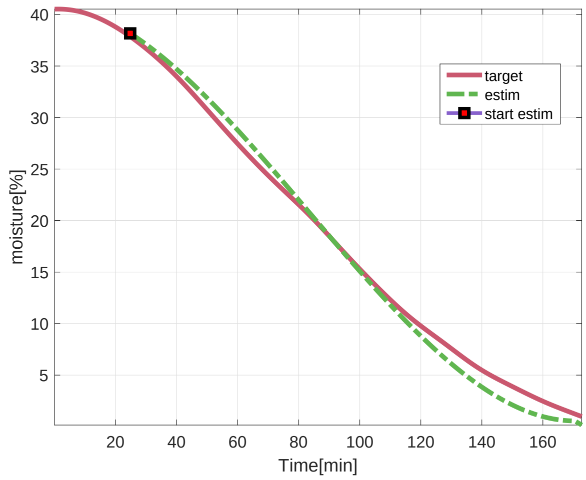


Fig. 6: An example of laundry moisture content estimation using the 3^{rd} degree polynomial model in time, on y-axis the moisture content in percentage [the square represents the time in which the estimation is provided during drying cycle]

polynomial in time which is smooth in contrast to results obtained with the other approach.

Table 2: 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 the a certain value of the error quantified using the RMSE metric defined before.

Methods	**performance target***	**low error variance**	**algorithm calibration**	**implementation facility**
Method 1		✓	✓	
Method 2	✓		✓	✓
Current algorithm	✓	✓		✓

Table 2 summarizes the comparison between (data-driven) methods proposed in this work. Method named Current algorithm[11] refers to the current procedure implemented online on machines and it is mentioned here in order to make clear

[11]Details about it will be omitted here because of intellectual property rights.

the features of each proposed procedure and the contribution for each item of interest[12].

The last two columns in Table 2 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, Method 1 suffers from an additional complexity level because model 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 method is simpler to represent online because the model is a time series with few parameters.
Method 2 has been chosen because it shows good results in terms of performance target and it makes the calibration process easier respect to the current algorithm once implemented on machine; the current focus is on the improvement of the error variance (due to the variability of estimations) using other different regression techniques.

6 Conclusion

This work dealt with the development of Soft Sensors to estimate laundry moisture content in household Heat Pump Washer-Dryers during drying in order to improve drying performance and user experience. This paper focuses on:

1. the use of different approaches to obtain the estimation of laundry moisture content during drying cycles in domestic washer-dryers;

2. the determination of the best model to provide the estimation of laundry moisture content as a time series obtaining a smooth curve to predict the End-of-the-Cycle automatically.

As regards the first point, experimental data were used to train models with available signals as predictors and laundry moisture content samples as the output in a supervised manner. Models were chosen to try the Symbolic Regression approach which exploits the Genetic Programming to construct the best model from data and provides coefficients for such a model.

The operating difference between the two proposed methods relies on the nature of the model: in Symbolic Regression predictors are signals and they are combined in (non-linear) ways, while using the 3^{rd} degree polynomial model (equation (3)) the prediction is a time series and the difficulty is on the online estimation of the parameters of such a model.
Main results of the work described in this paper can be summarized as below:

[12]Items 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.

- The best non-linear models selected by Symbolic Regression always involve main signals of temperatures and motor torque (or signals that are elaboration of those) which confirms that these signals are the most useful for our purpose; the models provided by this method use available signals as predictors which affect the estimations with their noise;

- the best model for the description of laundry moisture content as a time series is the 3^{rd} degree polynomial model (equation (3)) and it comes from a comparison between several options from literature. It turns out that this model is the best model to describe laundry moisture content during drying cycles basing on our available real data.

In addition, two strengths of this work needs to be underlined: (i) the novelty of the work: to the best of our knowledge there's no similar cases in literature concerning the online estimation of parameters of a moisture transfer model for fabric care applications; (ii) an effective 'soft' approach: no impact on hardware and no dedicated Design of Experiments (DOE) was required.

The research employed in this paper is a part of a doctoral thesis, which is still in development. Such doctoral thesis is expected to be completed in the coming months and is focused on the development of SS for fabric care major appliances with the purpose of improving the performances of the implemented drying algorithm and simplifying the calibration phases of them.

References

[1] Shang C, Yang F, Huang D, Lyu W (2014) Data-driven soft sensor development based on deep learning technique. Journal of Process Control 24(3):223–233

[2] Trevor H, Robert T, JH F (2009) The elements of statistical learning: data mining, inference, and prediction. Springer

[3] Kadlec P, Gabrys B, Strandt S (2009) Data-driven soft sensors in the process industry. Computers & Chemical Engineering 33(4):795–814

[4] Soares SG (2015) Ensemble learning methodologies for soft sensor development in industrial processes. PhD thesis, Departamento de Engenharia Electrotcnica e de Computadores da Faculdade de Cincias e Tecnologia da Universidade de Coimbra

[5] Susto GA, Zambonin G, Altinier F, Pesavento E, Beghi A (2018) A soft sensing approach for clothes load estimation in consumer washing machines. In: 2018 IEEE Conference on Control Technology and Applications (CCTA), IEEE, pp 1252–1257

[6] Susto GA, Maggipinto M, Zannon G, Altinier F, Pesavento E, Beghi A (2018) Machine learning-based laundry weight estimation for vertical axis washing machines. In: 2018 European Control Conference (ECC), IEEE, pp 3179–3184

[7] Zambonin G, Altinier F, Corso L, Sessolo M, Beghi A, Susto GA (2018) Soft sensors for estimating laundry weight in household heat pump tumble dryers. In: 2018 IEEE 14th International Conference on Automation Science and Engineering (CASE), IEEE, pp 774–779

[8] Susto GA, Vettore L, Zambonin G, Altinier F, Beninato D, Girotto T, Rampazzo M, Beghi A (2019) A machine learning-based soft sensor for laundry load fabric typology estimation in household washer-dryers. In: IFAC Conference on Intelligent Control and Automation Science (ICONS), Belfast

[9] Schölkopf B, Smola AJ, Bach F, et al (2002) Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press

[10] Scheidegger AE (1958) The physics of flow through porous media. Soil Science 86(6):355

[11] Luikov AV (1964) Heat and mass transfer in capillary-porous bodies. In: Advances in heat transfer, vol 1, Elsevier, pp 123–184

[12] Doymaz I (2004) Effect of dipping treatment on air drying of plums. Journal of Food Engineering 64(4):465–470

[13] Kucuk H, Midilli A, Kilic A, Dincer I (2014) A review on thin-layer drying-curve equations. Drying Technology 32(7):757–773

[14] Younis M, Abdelkarim D, El-Abdein AZ (2018) Kinetics and mathematical modeling of infrared thin-layer drying of garlic slices. Saudi Journal of Biological Sciences 25(2):332–338

[15] Dincer I (1998) Energy and environmental impacts: present and future perspectives. Energy Sources 20(4-5):427–453

[16] Ketelaars A, Pel L, Coumans W, Kerkhof P (1995) Drying kinetics: a comparison of diffusion coefficients from moisture concentration profiles and drying curves. Chemical Engineering science 50(7):1187–1191

[17] Sun SH, Marrero TR (1996) Experimental study of simultaneous heat and moisture transfer around single short porous cylinders during convection drying by a psychrometry method. International journal of heat and mass transfer 39(17):3559–3565

[18] Sahin A, Dincer I, Yilbas B, Hussain M (2002) Determination of drying times for regular multi-dimensional objects. International Journal of Heat and Mass Transfer 45(8):1757–1766

[19] Prasertsan S, Saen-Saby P (1998) Heat pump drying of agricultural materials. Drying technology 16(1-2):235–250

[20] Haghi A, Ghanadzadeh H (2005) A study of thermal drying process. Indian Journal of Chemical Technology 12(6):654–663

[21] Sander A, Kardum JP (2009) Experimental validation of thin-layer drying models. Chemical Engineering & Technology: Industrial Chemistry-Plant Equipment-Process Engineering-Biotechnology 32(4):590–599

[22] Hamdaoui M, Baffoun A, Chaaben KB, Hamdaoui F (2013) Experimental study and mathematical model to follow the drying phenomenon of knitted textile fabric. Journal of Engineered Fibers and Fabrics 8(3):155892501300800,308

[23] Dincer I, Zamfirescu C (2016) Drying phenomena: theory and applications. John Wiley & Sons

[24] Crank J, et al (1979) The mathematics of diffusion. Oxford university press

[25] Pan N, Gibson P (2006) Thermal and moisture transport in fibrous materials. Woodhead Publishing

[26] Luikov Av (2012) Analytical heat diffusion theory. Elsevier

[27] Sharma S, Tambe SS (2014) Soft-sensor development for biochemical systems using genetic programming. Biochemical Engineering Journal 85:89–100

[28] Orzechowski P, La Cava W, Moore JH (2018) Where are we now?: a large benchmark study of recent symbolic regression methods. In: Proceedings of the Genetic and Evolutionary Computation Conference, ACM, pp 1183–1190

[29] Searson DP (2015) Gptips 2: an open-source software platform for symbolic data mining. In: Handbook of genetic programming applications, Springer, pp 551–573