July 2018

Analysis of Data from the Industrial Machinery within the Hot Rolling Process for Predictive Maintenance

J.R. RUIZ-SARMIENTO [a,b,1], J. MONROY [a,b], F.A. MORENO [a,b], J.M. BONELO [c] and J. GONZALEZ-JIMENEZ [a,b]

[a] *Machine Perception and Intelligent Robotics Group*
System Engineering and Auto. Dept., University of Malaga
[b] *Institute of Biomedical Research of Malaga (IBIMA)*
[c] *ACERINOX Europa S.A.U.*

Abstract. Manufacturing industries are increasingly adopting data-driven, decision making systems towards the Industry 4.0 paradigm. In the context of this data revolution, the innovative SiMoDiM project aims at developing a smart predictive maintenance system for the stainless steel industry. In its first stage, it focuses on the assets within the hot rolling process, one of the core components involved in the manufacturing of steel sheets, and more specifically on the coiler drums of Steckel mills. These drums operate under mechanical and thermal stresses that degrade them, and their replacements directly impact the product valor chain. In this work we present the data analysis stage of SiMoDiM, where the huge amount of available historical and real-time data from the hot rolling process (collected by onboard sensors in the mills) are studied in order to find which variables and descriptors are valid indicators of the coiler drums' conditions. This analysis is the first step towards an intelligent system that takes advantage of such descriptions for performing a predictive maintenance of the machinery.

Keywords. Data Analysis, Predictive maintenance, Industry 4.0, Intelligent prognostics tools, Condition monitoring, E-maintenance, SiMoDiM project

1. Introduction

The recent changes in economic, social, and environmental requirements for the manufacturing industries, as well as the intensive monitoring of processes and systems within companies' assets producing huge amounts of rich data linked by connected networks, have promoted the emergence of the *Industry 4.0* paradigm [1,2,3]. The insight behind it is to evolve from control-based to smart factories, able to predict behaviours in customers, processes and systems, in order to anticipate them and self-adjust their operations at different levels [4,5].

A promising field of application of the Industry 4.0 paradigm is the reduction of costly, unscheduled downtime and unexpected breakdowns [6,7]. An important factor

[1] Corresponding Author; E-mail: jotaraul@uma.es

contributing to these breakdowns is the task of maintenance (that ultimately involves replacement of spare parts). Maintenance approaches turned from traditional *fail-and-fix* practices, which involve acting after the equipment fails, to *preventive* or *blindly proactive* ones, which assume a certain level of performance degradation (based on experience and human expertise) to carry out maintenance or replacing tasks. However, with the recent availability of data from processes and systems in a networked environment, it is possible to monitor the degradation of the machinery rather than detecting the faults and, ultimately, to optimize asset utilization in the facility. This approach is commonly referred as *predictive maintenance* or *e-maintenance* methodology [8]. It relies on the fact that machines usually go through a measurable process of degradation before they fail, hence enabling a prediction of when a preventive maintenance must be carried out. As a consequence, predictive systems lead to a reduction in costs, an increment of operation efficiency, and an improvement of the product quality [9].

Despite their clear advantages, predictive systems are not so common in real factories due to its challenging implementation [4,6]. Perhaps one of the most significant challenges is that, although those factories could incorporate modern monitoring tools, the real-time, produced data must be rendered in a usable form for its exploitation [10]. Information coming from logistics, scheduling, and production (through sensors mounted in the machinery) comprises a huge amount of data series that need to be summarized by analytics and modelling applications. This elaborated information would permit the manufacturers to gain awareness about the state of their systems and properly schedule maintenance operations.

In this work we present the initial steps towards the implementation of this predictive maintenance methodology in a stainless steel factory of *ACERINOX Europa S.A.U.* [11], one of the most competitive groups in the world in stainless steel production, a widely used product in manufacturing and construction. Concretely, we focus on one critical procedure in this industry, the multipass *Hot Rolling*, a mill process which involves rolling the steel at a high temperature, enabling an easy shaped of it. From the components involved in this process, the spotlight is on the drums within the coilers, crucial parts for the proper mill performance that operate by rolling the steel under mechanical and thermal stresses. The e-maintenance related to this rolling process is part of the challenging and innovative project *SiMoDiM*, which aims at developing a novel monitoring and diagnostic system for the stainless steel industry, seeking the digital transformation of the production processes in order to achieve intelligent, more efficient, competitive and flexible factories. A preliminary version of this work, completed by the one described here, was presented in [12].

Given the aforementioned huge amount of real-time and historical data to process (a *Big Data* problem [5]), the system for the predictive maintenance of the coilers' drums has to be designed in two phases: a first one where the historical data is analyzed in order to find a proper way to summarize them, and a second step where this summary is used to fit the predictive model (usually through a training-evaluation loop). There exist approaches that fuse these two phases, *i.e.* it is not necessary a previous step to resume the data since the own predictive technique does it internally. However, by doing so it is lost the access to the summarized data, which is a rich source of information that can be used to, for example, alleviate the heavy demand of memory for storage, or assess the system performance.

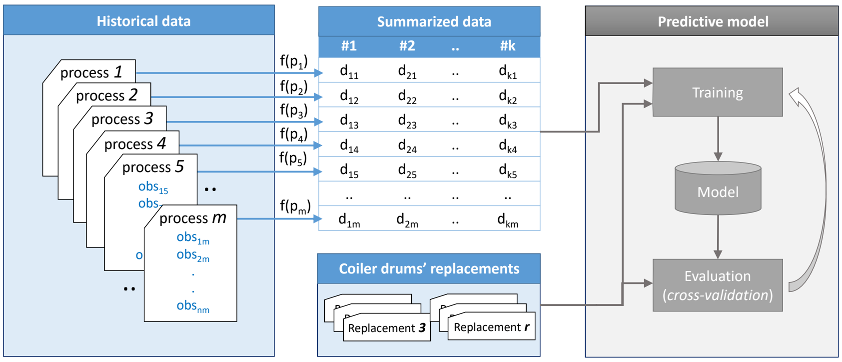

Figure 1. Information and processes involved in the development of a predictive maintenance system in the scope of the SiMoDiM project.

Fig. 1 illustrates the interactions and information involved in these two phases. This work focuses on the data analysis step, which encompasses the blue areas, while the fitting of the predictive model (gray ones) is left for future work. Our goal is to build a *procedure* or *function* that, taking as input the recorded or real-time information about a hot rolling process p_i, summarizes it through a vector of k descriptors $[d_{1i}, d_{2i}, \ldots, d_{ki}]$ [13]. This description has to be completed in such a way that the predictive system could retrieve from it the state of the monitored machinery, so it must be a degradation indicator. After a review of related work in the field (Section 2), and an outline of the hot rolling process (Section 3), we analyze the collected dataset (a vast set of data about the hot rolling process from the ACERINOX factory in Cadiz, Spain, see Section 4), and present the visual and numerical analysis carried out to design this summarization function (Section 5). The paper is concluded with a discussion about the obtained results and future work (see Section 6).

2. Related Work

The goal of the Industry 4.0 is to combine factories with modern technologies like Cyber-physical systems (CPS), Internet of Things (IoT), Internet of Services (IoS), or Big Data, pursuing modular and efficient manufacturing systems [1]. Essentially, CPS produce real-time information about their state, which is accessible by means of the IoT, and stored and analyzed through Big Data techniques. The result of such analysis can be visualized and further processed through the IoS, hence resulting in a manufacturing intelligence from real-time data that supports accurate and timely decision-making [5]. The number of works addressing the cooperation of these technologies in the manufacturing field is growing, as illustrated by the recent surveys by Liao et al. [2] or Sreedharan and Unnikrishnan [3]. The areas where the Industry 4.0 induces novel procedures are diverse, including, for example, process and planning (focusing on the reduction of waste and on the increase of product value), supply chain, transport and logistics, health and safety, product design, or the one addressed in this paper, maintenance and diagnosis.

In this context, Lee, Kao and Yang [6] proposed a framework for self-aware and self-maintained machines that can extract meaningful information from big data and perform an intelligent decision making regarding maintenance operations. A case of study is presented, where the target machinery was a heavy-duty equipment vehicle used in mining and construction. However, they do not provide information about how the produced data is summarized within the framework. Li, Wang and He [7] conducted a review of maintenance strategies, including corrective maintenance, preventive maintenance, and predictive maintenance, and investigated the potentials and trends towards the former in the scope of the Green Monitoring project. In this case it is superficially mentioned the domains where the data is described, including time, time-frequency, frequency, and wavelet domains, without further information. One of these authors, He, introduced in [14] other projects where the methodologies described in the former were also applied, like for example the WINDSENSE [15] or MonitorX [16] projects.

General predictive manufacturing systems were broadly discussed by Lee *et al.* [9], who presented a conceptual framework for their development and addressed the so-called *Prognostics and health management* (PHM) tool. PHM focuses on the estimation of the health of a production asset, the detection of incipient failures and prediction of the next fault event. This concept is closely related to the predictive maintenance one. They employ the Watchdog Agent [17], a toolkit for the prediction of performance of products and machines, as the component to conduct the PHM predictive analytics in a case study with a cutting tool. Again, it is not provided information about how the statistical summary of the variables of interest is extracted. Unlike the aforementioned works, in this paper we describe the resorted processes and techniques during the analysis of the available data in the scope of the SiMoDiM project, which encompasses information from: sensors installed in a hot rolling mill, configuration variables, as well as logistic parameters. This analysis is conducted to select a set of variables and descriptors to be leveraged by a predictive maintenance system, enabling, this way, an informed decision making about maintenance operations.

3. Steckel Hot Rolling Process

In metalworking, rolling is a metal forming process in which metal stock is passed through one or more pairs of rolls within a mill to reduce and uniform the thickness. Hot rolling refers to the case when the rolling process takes place employing metal that is preheated above its recrystallization temperature, typically over 1700 °F [18].

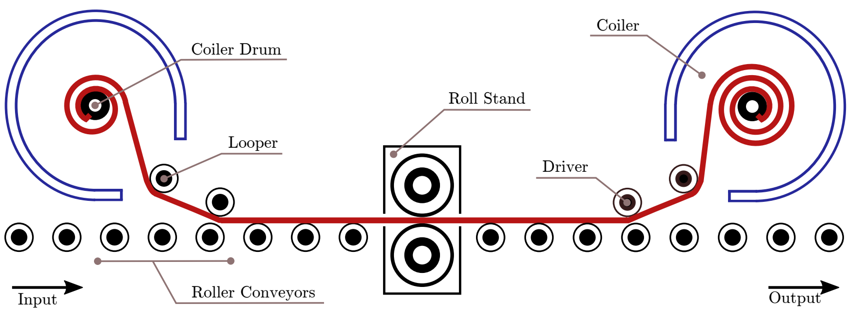

Figure 2. Scheme of the Steckel hot rolling mill. The stainless steel sheet (in red) is heated and worked in the mill through one or multiple passes, until desired thickness is obtained.

The starting material are large pieces of metal (like semi-finished casting products, often called plates), which, after being heated, are worked to reduce their thickness. For that, the material may need to pass one or several times through the mill, iteratively reducing the thickness and increasing its length. Fig. 2 shows a scheme of the Steckel hot rolling mill [19] employed by ACERINOX, where the steel sheets run along the roller conveyors to be worked in the roll stand. If more than one pass is necessary, the metal sheets are coiled around the drum, and the process is repeated in the inverse direction (left-to-right, right-to-left). In a Steckel mill with this configuration the number of passes is always odd.

Due to the high temperatures of the process (the coilers contain a furnace to keep the steel temperature high), degradation of the machinery is common and a proper maintenance plan is mandatory to avoid costly and long production downtimes. In this respect, the coiler drums are the parts of the mill that fastest degrades because of the high temperatures and the friction against the material being rolled, therefore being the spare part which replacement we want to predict before it breaks.

4. The Data

The available data span over the years 2013-2016. For each month within that period, a number of files are provided describing each hot rolling process carried out in the ACERINOX factory in Cadiz, Spain. These files contain the value of 18 different variables measuring the processes' state after every 0.5 meters of rolled steel, including, for example, steel densities, coiler temperature, engines power or pressure and forces in the roll stand. The variables are directly measured in the rolling mill by a number of sensors provided by the own control system of the machine, as well as from additional sensors gathered by a data acquisition card commanded by *LabView*. The Fig. 3 shows an excerpt of one of these files. The files have a different number of measurements (each one codi-

1	Campaign	ID	Pass	Meter	Gap	Velocity	Power	Leveling	Bending	Inptu tension	Output tension
2	20130100108	04B7VV	1	0,8	11,87	75	6960	0,9	128	0	0
3	20130100108	04B7VV	1	1,3	11,87	75	6960	0,9	128	0	0
4	20130100108	04B7VV	1	1,8	11,95	89	7703	1	117	0	0
5	20130100108	04B7VV	1	2,3	11,95	89	7703	1	117	0	0
6	20130100108	04B7VV	1	2,8	12,02	100	8300	-1	120	0	0
7	20130100108	04B7VV	1	3,3	12,02	100	8300	-1	120	0	0
8	20130100108	04B7VV	1	3,8	12,1	106	8793	-1	125	0	0
9	20130100108	04B7VV	1	4,3	12,1	106	8793	-1	125	0	0
10	20130100108	04B7VV	1	4,8	12,17	110	8999	-1	123	0	0
11	20130100108	04B7VV	1	5,3	12,17	110	8999	-1	123	0	0
12	20130100108	04B7VV	1	5,8	12,64	113	8811	-1	121	0	0
13	20130100108	04B7VV	1	6,3	12,64	113	8811	-1	121	0	0
14	20130100108	04B7VV	1	6,8	12,83	114	8670	-1	118	0	1,73
15	20130100108	04B7VV	1	7,3	12,83	114	8670	-1	118	0	1,73
16	20130100108	04B7VV	1	7,8	12,9	117	8718	-1	124	0	0,48
17	20130100108	04B7VV	1	8,3	12,9	117	8718	-1	124	0	0,48
18	20130100108	04B7VV	1	8,8	12,98	117	8596	-1	125	0	0,48
19	20130100108	04B7VV	1	9,3	12,98	117	8596	-1	125	0	0
20	20130100108	04B7VV	1	9,8	13,05	117	8672	-1	123	0	0
21	20130100108	04B7VV	1	10,3	13,05	117	8672	-1	123	0	19,76
22	20130100108	04B7VV	1	10,8	13,13	119	8201	-1	121	0	19,76
23	20130100108	04B7VV	1	11,3	13,13	119	8201	-1	121	0	18,96
24	20130100108	04B7VV	1	11,8	13,2	119	8228	-0,4	121	0	18,96
25	20130100108	04B7VV	1	12,3	13,2	119	8228	-0,4	121	0	21
26	20130100108	04B7VV	1	12,8	13,27	117	8204	-0,3	122	0	21

Figure 3. Excerpt of a file containing information relative to a hot rolling process. Note that the measurements are taken each 0.5 meters of steel sheet processed.

fied in a row), which depends on the number of passes of the steel sheet in the mill and the final thickness desired. In the case of the process in the figure, the steel completed 7 passes and the file has a total of 5,615 rows.

Additionally, each process is identified by two *logistic variables* (number of campaign and steel plate identification, the two first columns in Fig. 3). There are also provided 18 *meta-variables* or *configuration variables* about the process, which consist of 8 variables reporting the properties of the steel plate (*e.g.* steel type, weight, length and thickness at the entrance, etc.), and 10 more configuring the behaviour of the process itself (work code, number of passes in the mill, coilers' temperature etc.).

The resultant dataset is vast, containing a total of 118,484 hot rolling processes (recall that each process, in its turn, consists of thousands of measurements), divided into 7,351 with one pass, 6,523 of 3 passes, 65,704 processes with 5 passes, and 38,906 of 7 passes or more. This data, as well as the real-time information produced by each hot rolling process in the factory, need to be summarized for their suitable exploitation by the predictive maintenance model. The next section describes how to achieve this.

5. Data Analysis

The goal of the data analysis step is to find the variables and descriptors that best summarize the (huge) available data from the hot rolling process in a meaningful way. In this context, with *meaningful* we refer to information that provides indications about the degradation of the coiler drums. In other words, with this study we are pursuing the design of the function $f(p_i)$ that summarizes each process p_i through a vector of descriptors $\mathbf{d} = [d_{1i}, d_{2i}, \ldots, d_{ki}]$ that reflects such degradation (recall Fig. 1). Typically, this analysis consists of a preliminary visual study of the variables' behaviour through graphical representations, carried out in Section 5.1, as well as the exploration of the possible features or descriptors to summarize the data (like measures of center or spread), conducted in Section 5.2.

From the 18 variables describing each hot-rolling process, two of them refer to the process progress (current pass and meters of steel processed), and other twelve reflect values that are consequence of the process configuration (aperture of the roll stand, temperature within the coilers, power in the engines, etc.). The remaining four, although can be also influenced by the process configuration, refer to behaviours of the steel sheets and the coilers during the hot rolling processes that are susceptible to be affected by the drums' state, hence being candidates to reflect their condition. Thereby, the four candidate variables to be studied are: *input* and *output-tension*, which measure the traction forces in both coilers, *leveling* that indicates the slope of the sheet being processed, and *bending* that measures its curvature. After the conducted analysis, we obtained the final set of promising variables and descriptors, which will be exploited by the smart predictive maintenance system (recall Fig. 1).

The possible interactions between the promising variables and the configuration ones have been also examined (see Section 5.3). This study reports which configuration variables influence the measurements of the candidates so, for a reliable performance, they have to be considered in the predictive system in order to isolate the effect of the drums' state. The conducted analyses have been carried out employing powerful Python packages for data management and visualization like numpy [20], matplotlib [21], pandas [22], or seaborn, among others.

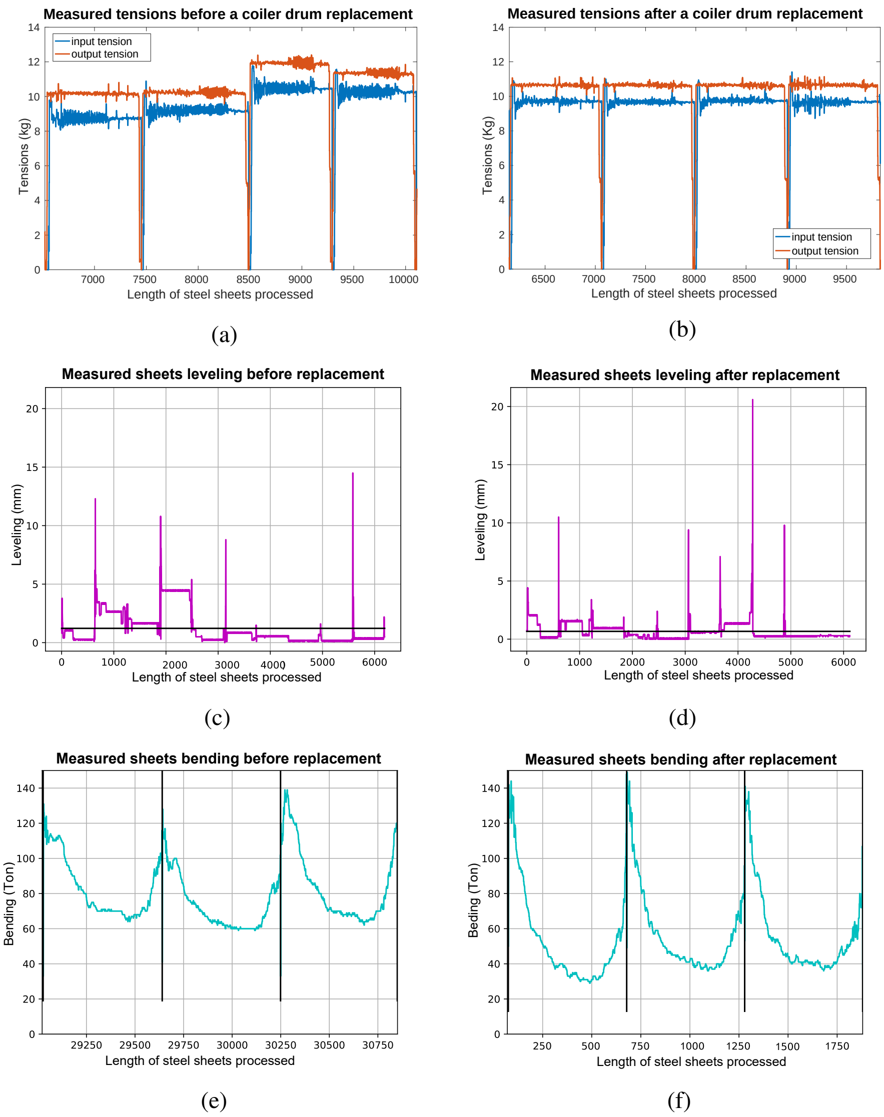

Figure 4. (a) and (b), measured tensions from the 4^{th} pass in four 5-passes processes just before (left) and after (right) a replacement of both coiler drums (vertical axes are kilograms). (c) and (d), measured sheets' leveling. (e) and (f), measured bending.

5.1. Visual analysis

A preliminary analysis of the data was visually performed aiming to figure out which descriptors could be the most appropriated to summarize the measurements from the four candidate variables, detect changes in their behaviour during the coiler drums lifetime, and identify possible relations with configuration variables. The experts from the factory pointed out that, in addition to other factors, the measurements of these variables are heavily influenced by the number of passes that the steel sheet does in the mill. To take this into account, we used that parameter to cluster the processes, being the resultant groups individually studied. This drastically reduced the data dispersion.

By means of this visual inspection it was also detected a strong correlation between the observed variables and the length of the sheets being processed. Fig. 4-(a) illustrates this for the case of the *input-tension* and *output-tension* variables, where we can clearly see that the shorter the sheet, the higher the tensions. This figure shows the input and output tensions measured during 4 processes (with sheets between 300 and 320 meters long) just before (a) and after (b) the replacement of both drums. According to the clustering done, in these processes the steel sheets accomplished 5 passes in the mill, and the figures correspond to the fourth pass. Exploring them, we can see how the tensions become more *stable* after the maintenance operation, which supports the expert's intuition that these variables can indeed be used to predict the drum's deterioration. This also incites to think that descriptors characterizing that stability could be useful.

The Fig. 4-(c) and (d) display the sheet leveling measurements for ten processes before and after a replacement, respectively. The black, horizontal line in those graphics represents the mean value of the measurements, and we can see how it decreases after the maintenance operation. The same occurs in the case of the bending measurements (see Fig. 4-(e) and (f)), where we can check that for three processes the *valley* values (those where the measurements are lower) decrease considerably, also decreasing its values on average.

5.2. Numerical analysis

This section outlines the numerical analysis performed to validate the previous insights and select promising variables and descriptors. For that, we performed a two-steps process: first, it was explored the descriptors' behaviours in the surroundings of four drum replacements, while in the second step we focused on the behaviour along the life-time of the coiler drums.

5.2.1. Analyzing variables' behaviour after a drum replacement

For this study we employed a window of 50 processes just after and before such replacements, selecting only processes with sheet lengths between 300 and 320 meters. From each data window, we computed a number of descriptors that summarize each process, and numerically analyzed the variability/discrepancies before and after the replacements. These descriptors are enumerated in the Table 1, and include among others: average and variance of each variable, oscillation value or FFT coefficients (analysis in the frecuency domain), etc.

In turn, the Table 2 reports an excerpt of the results obtained in this step. For example, in the case of the first studied replacement, the average values of the leveling and bending variables decreased considerably (*e.g.* from 1.01 to 0.82*mm.* and from 93.46 to 72.61*Ton.* respectively). They also oscillated less, which was checked by counting the number of times that these measures took values out of a certain range, reported as *hits* in the table. In the case of the tensions, in both cases the average values remained similar, however, considering the region when they reach a hill (*e.g.* around 10Kg. in Fig. 4-(a)), the standard deviations of their measures also decrease significantly: from 0.38 to 0.21 in the case of the input tension, and from 0.14 to 0.11 for the output one. Thereby, since the behaviour of the four (described) candidate variables differs before and after a maintenance operation, all of them passed this first verification/step.

Table 1. Computed descriptors for the different candidate variables (between parentheses, the number of total descriptors for each one). The descriptors of the output-tension variable have been omitted since they are the same as the input-tension ones. The check mark symbol (✓) identifies the promising descriptors chosen to feed the predictive system.

Leveling	(3)
mean	1
hits	1
hill *stdv*	1

Bending	(3)
mean	1
hits	1
minimum	1

Input Ten.	(12)		
mean	1	Max. amplitudes of FFT coefficients [✓]	3
stdv	1	Frequencies with max. amplitudes	3
hill mean	1	Skewness (frec. domain) [✓]	1
hill *stdv* [✓]	1	Kurtosis (frec. domain) [✓]	1

Table 2. Features extracted from the observed variables just 50 processes before and after three replacements of coiler drums.

		Replacement #1		Replacement #2		Replacement #3	
Variable	Descriptor	before	after	before	after	before	after
Leveling	mean	1.01	0.82	0.77	0.63	0.91	0.74
	hits	25	14	19	14	9	8
	hill *stdv*	0.21	0.18	0.14	0.11	0.15	0.10
Bending	mean	93.46	72.61	102.6	78.72	93.37	84.23
	hits	30	13	30	17	21	13
	minimum	84.86	60.12	91.20	65.26	81.08	70.06
Input tension	hill *stdv*	0.38	0.21	0.26	0.18	0.31	0.20
	Max. amplitude	0.069	0.038	0.052	0.029	0.055	0.031
Output tension	hill *stdv*	0.13	0.11	0.12	0.09	0.11	0.10
	Max. amplitude	0.021	0.016	0.017	0.013	0.017	0.013

5.2.2. Exploring variables' behaviour throughout drum life-time

In the second validation step, the behaviour of the described variables was checked through sequences of processes ranging from the instant of the installation of a fresh coiler drum until its replacement. The Fig. 5 illustrates this analysis for two variables: the sheet bending described through its mean values, and the coiler input tensions by its hill oscillations (standard deviations). In both cases it is shown their values for two sequences (as well as their trends), and we can check that although the behaviour of the input tensions' oscillations is similar (both increase similarly), it is not the case when analyzing the bending of the sheets.

To numerically drive the analysis, we computed the well-known coefficient of determination [23], denoted as R^2, which represents the portion of the variance in a dependent variable (in this case, the descriptions of the candidate variables) that is predictable from an independent variable (total number of hot-rolling processes completed). This is a good measure about in which magnitude a variable influences another one, that is, how the degradation of the coilers, which is clearly related to their usage, affects the behaviour of the candidate variables. To compute it, first, a linear regression model is built,

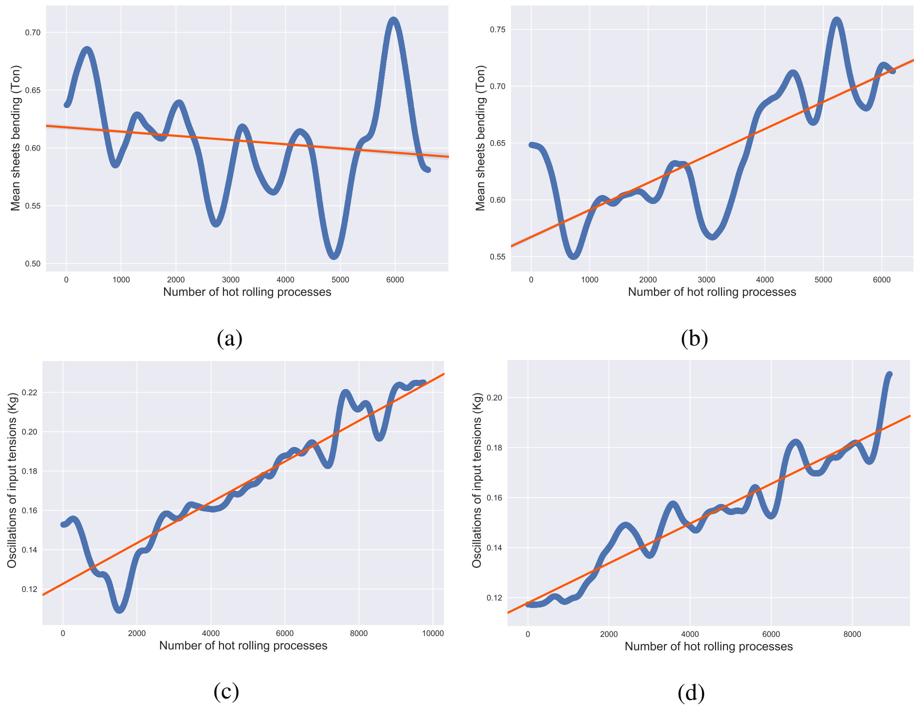

Figure 5. (a) and (b) illustrates the sheets' bending of the processes after the third and fourth drums' replacements, respectively, while (c) and (d) shows oscillations of input tensions for the processes after the same replacements. The data have been smoothed to improve their visualization.

and then the model residuals are analyzed. The output of this study was, for example, a coefficient of 0.34 for the oscillations of the input tension (*i.e.* the 34% of the oscillation can be predicted by the total number of hot-rolling processes completed), 0.26 for the oscillations of the output one, 0.01 for the mean bending values, and 0.11 for the mean leveling values[2].

The Table 1 shows a check mark next to the descriptors that reported a coefficient of determination above 0.2, comprising in this way the final set of promising variables and descriptors. Therefore, the vector of descriptors **d** describing each hot rolling process has 12 components (6 characterizing the input tensions and 6 for the output ones). These results also match the intuition of the experts from the factory, which argued that the sheets' bending and leveling can be strongly influenced by the state of other components of the rolling mill, and not just by the coiler drums.

5.3. Interaction of configuration-variables

For adjusting a reliable predictive system it is crucial to find and integrate the interactions between configuration-variables and candidate ones [9]. These interactions permit us to better model the behaviour of the studied variables, since their measurements would

[2]Notice that although the coefficient of determination for the mean leveling values is not negligible, it was discarded due to its irregular behaviour during the drums' lifetime.

Table 3. Coefficients of determination measuring the interaction of the configuration-variables with the oscillations of the input tension.

Meta-variable	R^2	Meta-variable	R^2	Meta-variable	R^2
STEEL_TYPE	0.15	WORK_CODE	0.01	FINISHING_CODE	0.15
REAL_WIDTH	0.20	REAL_LONG	0.00	REAL_WEIGHT	0.14
REAL_THICKNESS	0.00	NOMINAL_WIDTH	0.20	NOMINAL_WEIGHT	0.00
FINISH_TARGET_THICKNESS	0.20	FINISH_TARGET_WIDTH	0.18	DOWNLOAD_TEMP	0.00
NOM_FINISHER_THICKNESS	0.18	NUMBER_PASSES	0.01	PASS	0.25

probably not only depend on the coilers' conditions. Two of these interactions were discussed in Section 5.1: the number of passes of the steel sheet in the hot rolling process, and the length of the sheet within those passes. The goal here is to establish them in a principled way.

For that, the coefficients of determination between the candidate variables and the configuration ones were computed. The Table 3 reports these coefficients for the case of the oscillations of the input tensions (the results for the other descriptors were similar), and we can see how the strongest interaction is caused by the pass that the sheet is currently performing in the mill (first, second, third, etc.). This result matches the experts intuition. However, other configuration-variables also show high values for this coefficient, like the type of steel, the width of the sheet before the hot rolling process, or the target thickness after completing it, to name a few. This demonstrates the necessity of introducing these variables in the design step of the predictive model.

6. Conclusions

This work has described the data analysis step towards the development of a predictive maintenance system in the scope of the SiMoDiM project. The aim of SiMoDiM is the implementation of aspects from the Industry 4.0 in the production chain of stainless steel factories, concretely in the hot rolling process. As a case of study, it is explored the application of this revolutionary paradigm to the drums within the coilers in the Steckel hot rolling mills, critical components of such process. With the analysis carried out in this paper, a number of variables (collected from sensors installed in the Steckel mill) and descriptors have been selected to summarize in a usable way the huge amount of data produced by the mills monitoring. They have been also studied and detected interactions between these variables and those configuring the hot rolling processes, which must be taken into account by the predictive system.

Once the produced data is expressed in an usable form, our next milestone is to select and fit a suitable model for the predictive maintenance of the drums, which by being fed with such variables, allows us to detect the *optimum timing* for a preventive maintenance. At the moment we are obtaining promising results employing a Discrete Bayes Filter for that end. We also plan to add visual information to the model through the installation of cameras, aiming to detect possible deformations in the drums by means of computer vision techniques [24].

Acknowledgements: Work partially funded by the projects SiMoDiM ([OTRI-8.06/5.56.4826], [IC4-030000-2016-3]), and WISER ([DPI2014-55826-R]), financed by the Spanish Ministry of Industry, Energy and Tourism, and the Ministry of Economy, Industry and Competitiveness respectively.

References

[1] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann, "Industry 4.0," *Business & Information Systems Engineering*, vol. 6, no. 4, pp. 239–242, 2014.

[2] Y. Liao, F. Deschamps, E. de Freitas Rocha Loures, and L. F. P. Ramos, "Past, present and future of industry 4.0 - a systematic literature review and research agenda proposal," *International Journal of Production Research*, vol. 55, no. 12, pp. 3609–3629, 2017.

[3] R. Sreedharan.V and A. Unnikrishnan, "Moving towards industry 4.0: A systematic review," *International Journal of Production Research*, vol. 117, no. 20, pp. 929–936, 2017.

[4] B. Nikolic, J. Ignjatic, N. Suzic, B. Stevanov, and A. Rikalovic, "Predictive manufacturing systems in industry 4.0: Trends, benefits and challenges," in *Proceedings of the 28th DAAAM International Symposium*, 2017, pp. 796–802.

[5] P. ODonovan, K. Leahy, K. Bruton, and D. T. OSullivan, "Big data in manufacturing: a systematic mapping study," *Journal of Big Data*, vol. 2, no. 1, p. 20, 2015.

[6] J. Lee, H.-A. Kao, and S. Yang, "Service innovation and smart analytics for industry 4.0 and big data environment," *Procedia CIRP*, vol. 16, pp. 3–8, 2014.

[7] Z. Li, K. Wang, and Y. He, "Industry 4.0 – potentials for predictive maintenance," in *International Workshop of Advanced Manufacturing and Automation*, 2016, pp. 42–46.

[8] J. Lee, J. Ni, D. Djurdjanovic, H. Qiu, and H. Liao, "Intelligent prognostics tools and e-maintenance," *Computers in Industry*, vol. 57, no. 6, pp. 476–489, 2006.

[9] J. Lee, J. Lapira, S. Yang, and A. Kao, "Predictive manufacturing system - trends of next-generation production systems," in *11th IFAC Workshop on Intelligent Manufacturing Systems*, 2013, pp. 150–156.

[10] J. R. Ruiz-Sarmiento, C. Galindo, and J. Gonzalez-Jimenez, "Building multiversal semantic maps for mobile robot operation," *Knowledege-Based Systems*, vol. 119, pp. 257–272, mar 2017.

[11] "Acerinox Europe S.A.U. website," http://www.acerinox.com, 2018, [Online; accessed 20-07-2018].

[12] J. R. Ruiz-Sarmiento, J. Monroy, F.-A. Moreno, J. M. Bonelo, and J. Gonzalez-Jimenez, "Data analysis of the hot rolling process in a stainless steel factory for predictive maintenance," in *International Conference on Applications of Intelligent Systems (APPIS)*, 2018.

[13] J. R. Ruiz-Sarmiento, C. Galindo, and J. Gonzalez-Jimenez, "A survey on learning approaches for undirected graphical models. Application to scene object recognition," *International Journal of Approximate Reasoning*, vol. 83, pp. 434–451, apr 2017.

[14] K. Wang, "Intelligent predictive maintenance (ipdm) system – industry 4.0 scenario," *WIT Transactions on Engineering Sciences*, vol. 113, no. 1, pp. 259–268, 2016.

[15] K.-S. Wang, "Towards zero-defect manufacturing (zdm)—a data mining approach," *Advances in Manufacturing*, vol. 1, no. 1, pp. 62–74, Mar 2013.

[16] "MonitorX project," https://www.sintef.no/projectweb/monitorx/, 2018, [Online; accessed 20-07-2018].

[17] Z. Shi, J. Lee, and P. Cui, "Prognostics and health management solution development in labview: Watchdog agent® toolkit and case study," in *Prognostics and System Health Management Conference (PHM-Chengdu)*, Oct 2016, pp. 1–6.

[18] Metal Supermarkets, "Difference between hot rolled steel and cold rolled steel," https://goo.gl/PeoyKE, 2018, [Online; accessed 20-07-2018].

[19] Y. V. Konovalov and A. S. Khokhlov, "Benefits of steckel mills in rolling," *Steel in Translation*, vol. 43, no. 4, pp. 206–211, Apr 2013.

[20] T. E. Oliphant, *Guide to NumPy*, 2nd ed. USA: CreateSpace Independent Publishing Platform, 2015.

[21] J. D. Hunter, "Matplotlib: A 2d graphics environment," *Computing In Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[22] W. McKinney, "pandas: a foundational python library for data analysis and statistics," *Python for High Performance and Scientific Computing*, pp. 1–9, 2011.

[23] N. R. Draper and H. Smith, "Applied regression analysis," *Statistics in Medicine*, vol. 19, no. 22, pp. 3136–3139, 1998.

[24] J. R. Ruiz-Sarmiento, F.-A. Moreno, J. Monroy, and J. Gonzalez-Jimenez, "mVision, a toolbox for computer vision courses," in *12th annual International Technology, Education and Development Conference (INTED'18)*, mar 2018.