

# Advanced Monitoring of an Industrial Process integrating Several Sources of Information through a Data Warehouse

E. Sanz<sup>1,2,3</sup>, J.L. Matey<sup>1</sup>

<sup>1</sup>Paint Shop Maintenance  
SEAT, S.A.

Martorell, Barcelona, Spain

Email: [elma.sanz@seat.es](mailto:elma.sanz@seat.es), [esanz@iri.upc.edu](mailto:esanz@iri.upc.edu)

J. Blesa<sup>3</sup>, V. Puig<sup>2,3</sup>

<sup>2</sup>Automatic Control Department  
Technical University of Catalonia (UPC)

<sup>3</sup>Institut de Robòtica i Informàtica Industrial (CSIC-UPC)  
Barcelona, Spain

Email: {jblesa, vpuig}@iri.upc.edu

**Abstract**— This paper presents a methodology and architecture for the advanced monitoring of an industrial process integrating several sources of information using a data warehouse (DW) that include as metadata datamart to cross technical locations and equipments with the information given by the existing monitoring systems and the time dimension. The advanced monitoring includes functionalities that allow to diagnose faulty components and to prognose faulty situations when a problem occurs in the production process. A real car painting process is used for illustration purposes.

**Keywords**— *Advanced monitoring, Industry 4.0, Predictive Maintenance, system architecture, metadata, data sources, integration, Data Meaning, Data warehouse, data preparation, time dimension.*

## 1. Introduction

In the literature, the term Industry 4.0 is used to refer to a fourth industrial revolution with several concepts attached as e.g. Production 4.0, Predictive Maintenance, Internet of Things (IoT), IT + Automation, Operational Efficiency 2.0, Machine to Machine (M2M), Smart Factory, Additive Manufacturing, IT Security Management, Information and Communications Technologies (ICT), Cyber-Physical Production Systems (CPPSs).

Advanced Monitoring and Predictive Maintenance using Big Data Analytics in an industrial process are the focus of this paper. Predictive Maintenance is a maintenance methodology based on the principle that the machines actual degradation status is applied to decide the right time when a specific maintenance action is needed. The objective of this methodology is to reduce costs and increase availability (e.g.: before a fault occurs, the goal is to anticipate the occurrence time and apply a correcting it as quick as possible). Another important objective is to increase safety by preventing faults. This form of maintenance is made possible by the use of connected smart sensors on machines. These sensors measure conditions of the machine and provide this information to the monitoring system.

In everyday life, this kind of maintenance becomes more transparent. Manufacturing companies already apply several kinds of predictive maintenance for many years. To check the machinery condition, most companies have a program for using sensor measurements of the bearings vibration of rotating equipment, carry out analyses of the oil used in machinery and measuring temperature of critical electronic components. Based on the outcome of these data, it is decided whether maintenance is needed and, when it should be carried out.

Some common problems are missing data (data gaps) due to communication problems (e.g. no communication signal or failures), and problems related to the sensors (e.g. noise and peaks) due to malfunctions (e.g. sensor not calibrated or degraded equipment).

This intrinsic uncertainty in the raw data must be bounded somehow and reduced as much as possible in order to provide reliable data. This is achieved by applying some methodology to validate and reconstruct invalid data [1].

Nowadays, the need for processing large volumes of data to optimize process and maintenance management in the installations has increased exponentially. Most companies still do not analyse the process data that are collected in various shop floor systems (MES, SCADA) in an integrated way and with advanced data analytics algorithms. This process data contains “hidden” information about the status of the different machinery. However, this information is not easy to filter because of the huge amount of data, various sources where it is often stored and the data heterogeneity problem, for the semantic description of data (Data Meaning).

This paper is focused on development of a project called BiDrac that aims to study how this digital transformation and the opportunities that Industry 4.0 offers can be used to improve the process monitoring and installations maintenance of a particular industrial process: the Paint Shop of the car manufacturer Seat, S.A. in Martorell (Spain). The development of the proposed methodology is organised in a set of steps in order to adopt advanced techniques. This incremental manner allows the company to have a working

solution since the first stage, without requiring waiting the finalization of the last phase. The proposed approach has been designed to ease landing in Business Intelligence, and to be realized taking into account the IT standards and tools available in the company. This approach involves covering the following phases:

1. **From Raw Data to Database** considering data acquisition, data heterogeneity and data volume.
2. **Data Validation** to check the consistency and bounding the uncertainty of the raw data.
3. **Data Analytics** to extract relevant features and latent factors from noisy data.
4. **Knowledge Discovery** to provide a higher layer to identify (possible hidden) patterns and relationships inside the data.
5. **Application** to the current monitoring systems at the Seat Paint Shop with the aim to deploy all the mechanisms defined in the previous stages on BiDrac project in order to integrate all monitoring systems providing a robust and modular platform that can be run on-line.

The main objective of this work is to provide the Paint Shop with the suitable technologies and tools that improve the production and maintenance management and, in turn, that let them to anticipate problems that may arise.

The structure of the remainder of the papers is as follows: Section II presents the problem statement. Section III presents the proposed solution. Section IV illustrates the application of the proposed approach in the Paint Shop. Finally, Section V draws the main conclusions.

## II. Problem Statement

### A. Paint Shop Process

In the Paint Shop Process, during painting operations, metal vehicle bodies are initially put together, washed to remove the materials (including oily materials) used to protect body parts against corrosion and damage during stamping and transport, and treated with phosphate agents to prepare the surface for coating. The cleaned and surface-treated bodies are placed in a dip tank where they are electrolytically coated with paint polymers, followed by a series of spraying operations starting with primer line, base coat and finishing with clear coat. The painted bodies are baked in an oven to cure the coated paint materials.

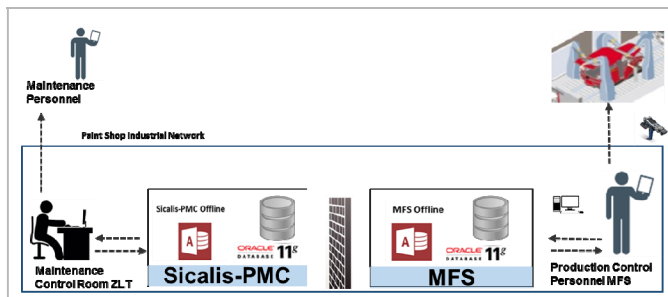


Figure 1: Monitoring systems in the Paint Shop.

### B. Current Monitoring System

In the Paint Shop Maintenance Department, there exist two real-time monitoring systems without connection between them (see Figure 1).

The Sicalis-PMC system monitors about 69.000 signals of different types (digital and analog). Over 5800 of these signals are stored in the server in a daily database with applying a one-minute resolution. This daily database contains more than 46 million data measurements offline. There is no connection between the real-time operational database (data from the last 60 layers) and the offline database (historical data) which makes impossible a continuous analysis of the signals. Moreover, there is a lack of suitable tools to present the values of signals in graphical form making more difficult the transformation of the data into information.

The other system is named MFS and monitors the production flow inside the process layout of the Paint Shop, the production values and products characteristics. The Paint Shop coats 2.400 bodies/day including three vehicle models (Ibiza, León and Q3). These vehicle models have 14 variants in total with different surfaces and they can be coated using about 24 different colours in three coating cabins (DL1, DL2, DL3). Figure 2 summarises the information gathered by each monitoring system.

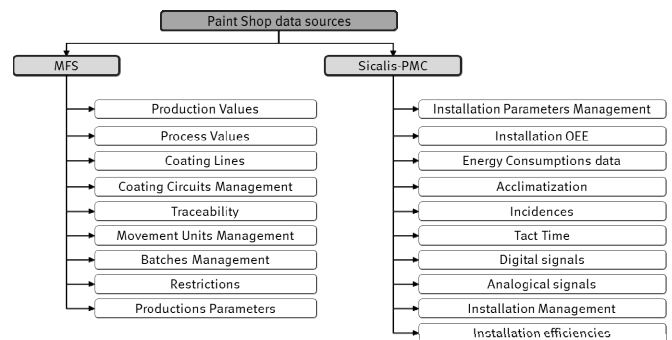


Figure 2: Current monitoring systems architecture scheme.

The current analysis tools consist of specific program codes developed for a particular analysis. The analysis process is complex, slow, and not very flexible. The current tools does not enable analysing correlated signals from different systems for establishing relations and extracting patterns at a given moment or in a period of time.

### C. Advanced Monitoring System Specifications

The advanced monitoring systems to be developed should be able to:

- Link the existing monitoring systems in the Paint Shop by designing and implementing a new Data Warehouse with an architecture that facilitates data analytics by crossing data from the different data bases.
- Integrate online/offline Data Analytics allowing the traceability of production and installation parameters (such as temperature, humidity or pressure).

- Implement an Early Warning System for operational problems
- Extract behavioural patterns in Paint Shop installations by analysing historical and real-time data.
- Correlate data from different variables to perform predictive analysis and deviations (errors in a system, locks, losses, etc.).
- Create and configure reports and dashboards by each analyst to have the required information at the moment that it is needed.
- Reduce costs by analysing and optimizing the use of the installations saving energy costs, spare parts or improve production (as e.g. reworks or strikes).
- Facilitate the application to other databases and allow simulations.

### III. Proposed Solution

#### A. Introduction

The implementation of the advanced monitoring system specified in previous section is based on development the BiDrac methodology that will involve the application of the Business Intelligence (BI) technologies to the organization and monitoring systems in the Maintenance Department of the Paint Shop at SEAT factory in Martorell.

BiDrac project is focused on reducing the large amount of data analytics time and resources needed nowadays in solving an installation failure, prevent installation failures and improving the installation efficiencies. The developed methodologies and tools will aim to detect parameter deviations or variables correlations in the Paint Shop Installation that would be able to detect incipient fault allowing actuating in advance and avoiding to stop the coating process. Also, it will allow increasing the installation efficiencies by studying the different technologies and making installation comparisons.

In case that one incipient fault is appearing in the system, a real-time monitoring of the signals will allow to study the evolution and to anticipate actions before the failure occurs. This is a big improvement for the maintenance analysts. Moreover, in case one failure occurs in one element of the process chain, the proposed methodology will allow to predict its effects and propagations to subsequent process components in order to predict its consequences in the future (blocking times for example) and to avoid/minimize them (minimizing chain effects).

#### B. BiDrac methodology development

The first step in the development of the BiDrac methodology involves studying the monitoring systems Sicalis-PMC and MFS, by analysing their functionalities and their data tables' architecture. A comparison of the data tables related in the documentation and the real data tables in the systems servers was carried out.

Then, a relational model involving all the data tables from the two monitoring systems and with a time dimension to join them was created. This has allowed designing the BiDrac Data

Warehouse with the BiDrac\_Master Datamart as a metadata repository (Figure 3). BiDrac is joining also the Online Mode with the Offline Mode Data Bases in its Data Warehouse [6][7].

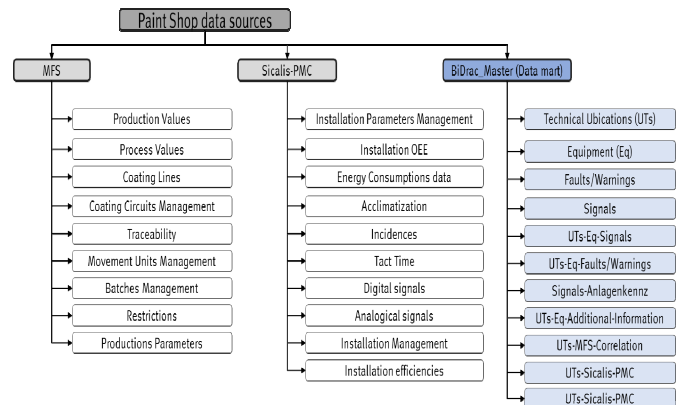


Figure 3: Proposed advanced monitoring architecture scheme.

The BiDrac Data Warehouse is going to contain cleaned, validated and prepared data for analysis. First, we feed it with extraction of online data using an extract, transform and load tool (ETL). Once the Data Warehouse receives data from the Online Mode, the ETL process for the Offline Data (that have recorded data from the systems during the last years) will be executed.

The analyses of the daily extractions of Sicalis-PMC and the weekly extractions of MFS were also obtained.

Figure 4 shows the BiDrac\_Master architecture scheme. As it can be seen in this figure, the BiDrac\_Master contains the mapping of the signals, parameters, warnings and alarms related to the Technical Ubications and the Equipments in the Paint Shop.

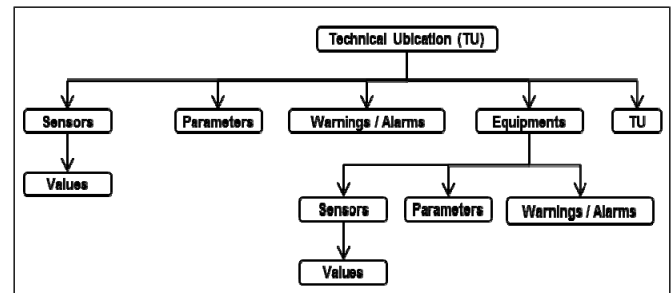


Figure 4: BiDrac\_Master architecture scheme.

Once the BiDrac Data Warehouse is available, using the SAP Business Object 4.1 Analytics, we design a Universe that helps to use this relational model to create the Reporting and the Analytics of BiDrac solution (see Figure 5).

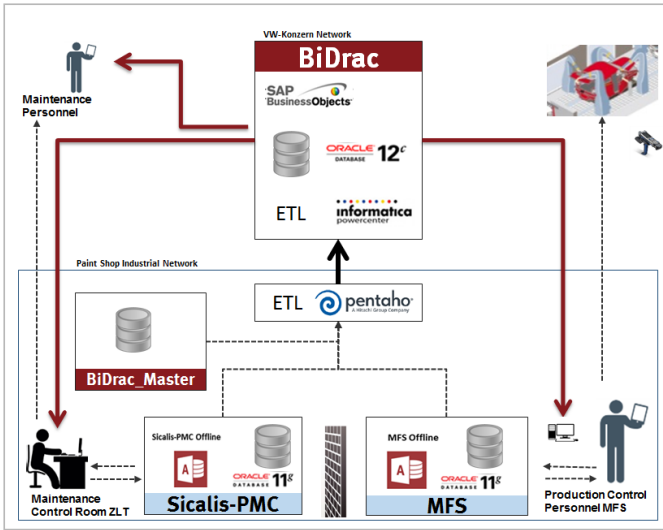


Figure 5: The BiDrac solution

## IV. Application Results

### A. Sensor data validation module

To illustrate the application of the BiDrac solution to the Paint Shop process, first we will present the application of the data validation methodologies to the signals involved in the production flow of Workshop 4 (T4) to Primer line in Workshop 5 (T5) described in Figure 6.

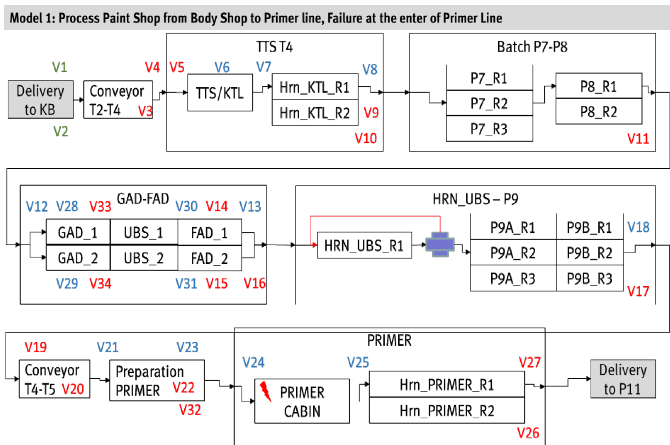


Figure 6: Model for the Production Flow analyses from T4 to Primer line in T5.

The models are obtained for each variable using historical data. Thus, each value can be compared against the model estimation to validate or invalidate the data [4].

Relating to the coating process described in section II.A, we choose 34 variables from 3 control systems (the two monitoring systems in the Paint Shop and the *Fabrik Information System* (FIS) that is the production scheduling system used in Seat, S.A.). Among the 34 variables, 18 variables come from the Sicalis-PMC Control System and are

computed using 78 different signals of this system. On the other hand, 14 variables come from the MFS Control System and 2 from the FIS System.

With the 34 variables and basic flow relations [2], 21 analytical redundancy relations (ARR) can be derived. From every ARR, a residual that should be zero can be computed. The list of variables and ARRs is detailed in Figure 7.

| ID_Rel | Form_Rel                       | System  |
|--------|--------------------------------|---------|
| R1     | $V1(k)=V2(k)$                  | FIS     |
| R2     | $V3(k)=V1(k)-V4(k)+V3(1)$      | FIS-PMC |
| R3     | $V6(k)=V5(k-T_1)$              | MFS-PMC |
| R4     | $V7(k)=V6(k-T_2)$              | MFS     |
| R5     | $V8(k)=V7(k-T_3)$              | MFS     |
| R6     | $V10(k)=V5(k)-V9(k)+V10(1)$    | MFS-PMC |
| R7     | $V10(k)=V5(k)-V8(k)+V10(1)$    | MFS-PMC |
| R8     | $V11(k)=V8(k)-V12(k)+V11(1)$   | MFS-PMC |
| R9     | $V14(k)=V28(k)-V30(k)+V14(1)$  | MFS-PMC |
| R10    | $V15(k)=V29(k)-V31(k)+V15(1)$  | MFS-PMC |
| R11    | $V13(k)=V16(k)$                | MFS-PMC |
| R12    | $V17(k)=V138(k)-V18(k)+V17(1)$ | MFS-PMC |
| R13    | $V19(k)=V18(k-T_4)$            | MFS-PMC |
| R14    | $V20(k)=V19(k)-V21(k)+V20(1)$  | MFS-PMC |
| R15    | $V22(k)=V21(k)-V23(k)+V22(1)$  | MFS-PMC |
| R16    | $V23(k)=V21(k-T_5)$            | MFS     |
| R17    | $V24(k)=V23(k-T_6)$            | MFS     |
| R18    | $V25(k)=V24(k-T_7)$            | MFS     |
| R19    | $V27(k)=V25(k-T_8)$            | MFS-PMC |
| R20    | $V26(k)=V24(k)-V27(k)+V26(1)$  | MFS-PMC |
| R21    | $V4(k)=V5(k)$                  | PMC     |

Figure 7: Production model ARR's

With the ARRs, we can build the Fault Signature Matrix (FSM) where for ARR (row) is indicated which variables (columns) are involved with one '1' in the crossing cells (see Figure 8). The FSM and the residuals computed using these relations help us to determine if exist any incongruence between the variables that are computed with the signals of the different data systems. If some residual/residuals is/are different from zero an inconsistency is detected. Then with the FSM and applying Artificial Intelligence consistency-based diagnosis techniques [5], we can isolate the signal or signals that are inconsistent with the relations. In practice, in order to avoid false alarms, a threshold for every residual is calibrated. If the residual value is smaller or equal to this threshold, the analytical redundancy is considered consistent and the contrary if the value of the residual is bigger than the threshold.

The information about the proposed diagnosis scheme described before can be used to anticipate the sensors maintenance operations.



|     | V1 | V2 | V3 | V4 | V5 | V6 | V7 | V8 | V9 | V10 | V11 | V12 | V13 | V14 | V15 | V16 | V17 | V18 | V19 | V20 | V21 | V22 | V23 | V24 | V25 | V26 | V27 | V28 | V29 | V30 | V31 |
|-----|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| R1  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R2  | 1  | 0  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R3  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R4  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R5  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R6  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 1  | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R7  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 1  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R8  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R9  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 0   | 1   | 0   |
| R10 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 0   | 1   |
| R11 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 1   | 0   | 0   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R12 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 1   | 0   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R13 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R14 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R15 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R16 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R17 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| R18 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| R19 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 1   | 0   | 0   | 0   | 0   | 0   | 0   |
| R20 | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 1   | 0   | 1   | 0   | 0   | 0   | 0   | 0   |
| R21 | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |

Figure 8: Production model FSM

As an example, we show residuals R6 and R7 in Figure 9 that are derived from the subtraction of the occupancy measured in the TTS\_T4 process and the other input and output volumes signals available in the monitoring systems (PMC and MFS). From the residuals, we can clearly see that the R7 shows coherence between the signals and, instead, the residual R6 shows incoherence. If we look at these relations in the FSM, we can determine that the problem is the signal V9 corresponding to the output volume of the TTS T4 PMC.

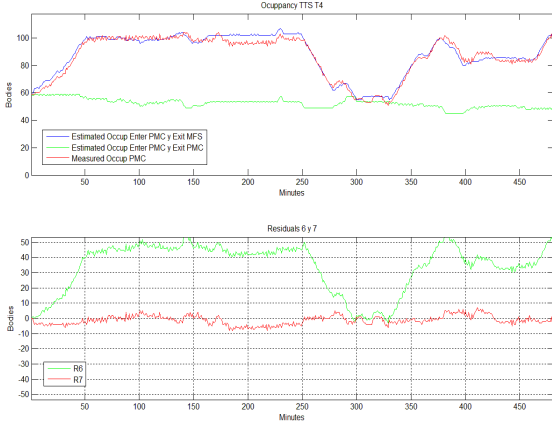


Figure 9: Coherence Tests in the Occupation of TTS T4 and its residuals R6 and R7.

The use of this methodology guarantees the quality and consistency of the signals received from the two monitoring systems and the increase the model's robustness in the presence of sensor faults (e.g.: if an occupation signal provides an incoherent value, this value can be replaced by its estimation from the input and output volumes of the same process).

### B. Prognosis module

Once the sensor values are validated, we can extract conclusions regarding the volumes and occupancies in the different processes and identify potential problems (e.g.: breakdowns, blockages or body absences).

Figure 10 shows an example of the effect of one robot failure in the Primer Cabin. It can be observed the increase of occupation of the Primer process and the rest of subsequent processes occupations upstream of the UBS oven. Paint drying is a critical process because of the time spent by the

body inside the oven may not exceed the maximum allowed heating time (25 min) and temperature.

To illustrate the forecasting capability of the prognosis module used in the proposed approach, we will evaluate the Remaining Useful Life (RUL) based on a predetermined Failure Threshold (FT) based on the maximum occupancy. As proposed in [3], the RUL is given by:

$$RUL \in N | \hat{y}(t + RUL | t) = FT \quad (1)$$

where  $\hat{y}(t + RUL | t)$  is the RUL-step ahead forecast at time  $t$  of a given predictive model ( $\hat{y}$ ).

The Brown's Double Exponential Smoothing is used to derive the predictive models from the data collected [8]. It is a multi-step forecast that can be expressed as follows:

$$y_1(t) = \alpha y(t) + (1 - \alpha) y_1(t - 1) \quad (2)$$

$$y_2(t) = \alpha y_1(t) + (1 - \alpha) y_2(t - 1) \quad (3)$$

$$\hat{y}(t + h | t) = \left( 2 + \alpha \frac{h}{1 - \alpha} \right) y_1(t) - \left( 1 + \alpha \frac{h}{1 - \alpha} \right) y_2(t) \quad (4)$$

where  $h$  is the forecast horizon and  $\alpha$  the smoothing parameter. The  $\alpha$  parameter should be adjusted using historical data to obtain a good prediction result using a non-linear least squares approach.

The robot failure in the Primer Cabin occurs at time  $t = 250$  minutes in the first station of the Primer Cabin (e.g. failure of a robot) and there is a drop in the occupancy, i.e., in the number of bodies inside the Primer Process. Consequently, this causes the increase in the occupancy of processes and batches upstream. The occupation of the Preparation Primer shows no change because their occupancy was already saturated. The T4-T5 Bridge saturates quickly because it is very close to saturation and the P9 batch starts increasing its occupancy.

With the pending drop in Primer occupancy and the observed pending rise of the P9 batch occupancy an anomaly at the entrance (block) of the Primer process is detected. Using the proposed prognosis technique, we can estimate the RUL that in this case provides the time of reaching the maximum capacity of P9 batch, preserving the maximum capacity of the UBS Oven. This ensures the evacuation of the bodies inside the UBS Oven and avoids blocking the exit of the UBS Oven with vehicles inside the process.

In five minutes, it is determined that this pending drop in Primer occupancy corresponding to a problem at the entrance of vehicles in the process. After the problem is detected, the occupancy forecasting based on (2)-(4) with smoothing parameter  $\alpha = 0.8$  is used to use to predict the occupancy in P9. Then, we use this forecast to compute the RUL by means of (1) considering FT equal to the maximum occupancy of P9 batch minus the maximum occupancy of the UBS Oven to ensure the UBS Oven is not being to be blocked. The computed RUL is equal to 68 minutes. This means that if in this time we cannot solve the failure of the robot in the Primer

Cabin, we cannot guarantee that the UBS Oven is not being to be blocked.

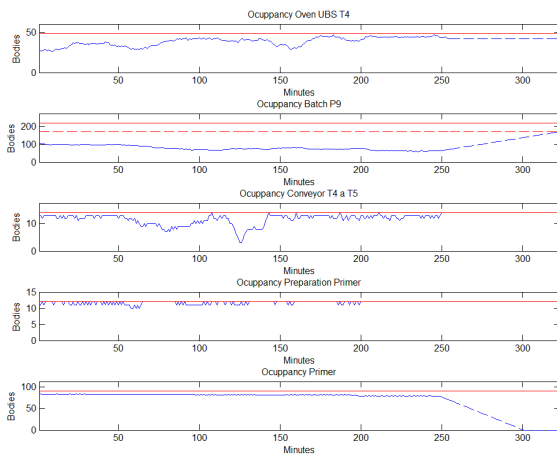


Figure 10: Effects in the occupancy of the processes when a robot failure is present in the Primer Cabin at  $t=250$  minutes.

## v. Conclusions

In this paper, an advanced monitoring system for the Paint Shop Maintenance Department of an automotive factory has been proposed. The proposed solution integrates two data sources used for the different monitoring systems and a new metadata datamart in a Data Warehouse. Thanks to this integration the advanced monitoring and predictive maintenance of the considered industrial process was possible to be implemented.

The proposed solution adds in the architecture scheme a new datamart with metadata related to the “Things” of the “Internet of Things (IoT)” concept: Technical Ubications and Equipments in the Paint Shop including all its signals, parameters, warnings, alarms, etc. This new architecture is focused on the data preparation process in order to choose those data that are significant for the advanced monitoring and predictive maintenance techniques. The data from the two sources vary in time. So in order to check them, the proposed solution adds also a time dimension in the metadata datamart.

For illustration purposes, the application results of the sensor data validation and prognosis modules are illustrated using realistic scenarios in the production flow of the car painting process. As future work, more advanced data analytics algorithms are being implemented in the advanced monitoring and maintenance system are being deployed to further exploit the information available allowing to carry out more sophisticated analyses that the ones currently implemented.

## Acknowledgment

We have to express our appreciation to all the Departments in the Paint Shop and in the SEAT, S.A. who provided insight and expertise that greatly assisted the research. Special thanks to all the workers in the Maintenance Department for their contributions, collaborations and the fieldworks we did in the installations.

This work has been partially funded by the Spanish Government (MINECO) through the project CICYT ECOCIS (ref.DPI2013-48243- C2-1-R), by MINECO and FEDER through the project CICYT HARCRICS (ref.DPI2014-58104-R).

## References

- [1] M. Jordan. *Frontiers in massive data analysis*. National Academies Press, 2013.
- [2] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki, *Diagnosis and fault-tolerant control*. Springer-Verlag Berlin Heidelberg, 2006.
- [3] T. Escobet, J. Quevedo, and V. Puig, “A Fault / Anomaly System Prognosis using a Data- driven Approach considering Uncertainty,” *IEEE World Congress on Computational Intelligence*, pp. 10–15, 2012.
- [4] D. Garcia, R. Creus, M. Minoves, X. Pardo, J. Quevedo, and V. Puig, “Prognosis of quality sensors in the Barcelona drinking water network”. *Conference on Control and Fault-Tolerant Systems, SysTol*, pp. 446-451. 2016.
- [5] R. Reiter. “A theory of diagnosis from the first principles”. *Artificial Intelligence*, 32, pp. 57-95. 1987.
- [6] R. Kimball and M. Ross. *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, Third Edition. John Wiley & Sons, 2013.
- [7] R. Kimball, *Data Warehouse Toolkit Classics: The Data Warehouse Toolkit*, 2nd Edition; *The Data Warehouse Lifecycle Toolkit*, 2nd Edition; *The Data Warehouse ETL Toolkit*. John Wiley & Sons, 2009.
- [8] R.G. Brown. *Smoothing Forecasting and Prediction of Discrete Time Series*. Englewood Cliffs, NJ: Prentice-Hall, 1963.