

### WORKSHOP 05

Digital Twins for CBM Apps. A Case Study in High-Speed Train Bearings Dt.

Prof. Adolfo Crespo Marquez Full Professor, University of Seville Spain

#### An Initiative by



Organized by



### Abstract

- The popularity of the concept of Digital Twins (DT) is emerging.
- DTs aim to replicate physical equipment and systems in the digital world through effective integration of data, models, and decision-support systems, promising a step change in productivity and sustainable performance.
- Several challenges remain to be addressed: General lack of conceptual basis, functional description, and a clear absence of fully established requirements.
- This workshop offers a generic framework for the functional description of a DT designed for intelligent maintenance purposes, and besides that, list a set of requirements features to fulfill when developing these tools, according to relevant scientific literature.
- The framework for the DT functional description and the DT requirements fulfillment has been tested in real CBM Applications of TALGO, a well-known high speed train manufacturer. Thanks to this exigent work environment, the methodology is sufficiently robust to be replicated in other operational contexts.





### Abstract

- Introduction to:
  - $\circ$  The problem
  - The Digital Maintenance Management Framework (DMM Framework)
  - **o** The Digital Twins (DT) and their Requirements
- Case Study: Train Axle Bearing CBM DT
  - CBM DT explanation using the DMM
  - Anomalies Detection for CBM
  - Failure Mode Classification for CBM
  - Data Analytics for Prognostics for CBM
  - $\circ$  Interaction with the CBM DT
  - Fullfilment of DT Requirements



### Introduction

Rapid advances in digital technologies, data analytics and artificial intelligence applied to maintenance.

These approaches have the potential to transform the way maintenance is managed, generating a deeper understanding of how complex industrial systems behave and perform, thus enabling us to manage them better.

In this context, data plays a pivotal role to enhance maintenance management processes. Data can now be extracted, prepared, and recorded, for specific decisionmaking maintenance processes, automatically (this is named ETL extraction-transformation-and-load of data). Then, intelligent assets management systems apps (IAMS Apps) support the different decision-making processes organizing the collection and the analysis of data.





### Introduction

IAMS Apps may also interact with additional tools such as simulation tools, providing extra analytical services, and they may add complementary data to the database records with results provided by these software elements. In addition to this asset knowledge discovering, creation and storing (Marquez et al., 2020), these IAMS Apps are provided by vendors together with business intelligence features or Apps (BI Apps).

The BI App is designed for the interaction with the end user and extract database records to present the information according to the reporting needs and end user requirements, on demand or at the time needed by the business. A simple data flow of the process is presented in Figure 1 (adapted from (Marquez et al., 2020).





### Introduction. Digital MM Framework



D Springer



## **Digital Twins and their Requirements**

Maintenance is starting to be the most common application of DTs, followed by Prognostic Health Management (PHM) and lifecycle optimization and the sectors where it is most applied are manufacturing, energy industry and aerospace (Errandonea, Beltrán and Arrizabalaga, 2020).

Some authors sustain that in order to be considered a Digital Twin, a model must have some specific characteristics:



Applicable to more equipment or failure
Combinable with different models and data
Extendable with new models development
Precise tracking of system behavior and status

Decision making and human interaction

General process and system architecture integration



### **Case Study.** The Train Axle Bearing CBM DT

In the broadest understanding, CBM include solutions detection, diagnostics, and prediction of failure modes that can be interpreted to maintenance decisionprovide making (Guillén, Crespo, Gómez, & Sanz, 2016). This case study DT has been elaborated to detect, classify, and predict train axle bearing failures using bearings monitored variables, in this case each bearing was only monitored capturing its temperature.



### **DT Explanation using the DMM Framework**



**OMAINTE** 20 VEARS



### **Anomalies Detection for CBM**

A train axle bearing temperature depends on a set of factors when the train is running at the uninterrupted regime: the type and dimensions of bearings, the antifrictional and hydrodynamic properties of the lubricant, the spaces between the bearing rollers and rings, the static and dynamical loads of the bearing, the train running speed, the duration of travel without stops, the ambient air temperature, and the road curves [(Lunys, Dailydka, & Bureika, 2015), (Mironov AA, 2008)] (see Figure 4).



Figure 4. Factors (physical model inputs) conditioning a train axle bearing temperature.





### **Anomalies Detection for CBM**

This workshop DT departs from the fact that the theoretical physical model to calculate axle bearing temperatures could be replaced by a data-driven bearing temperature model as in (Crespo Márquez, de la Fuente Carmona, Marcos, & Navarro, 2020).

The data-driven model inputs and outputs are presented in this Figure 5.

To estimate an axle bearing temperature, the remaining axle bearings temperatures plus the ambient temperature are the only inputs considered. This is the capital principle, and very innovative approach, to build all DT required predictive analytics.



Figure 5. Crespo et al. (2020) approach to predict axle bearing temperatures.





### **Anomalies Detection for CBM**

The anomalies detection rule designed could identify damaged bearings with 100% precision, at any speed of the train, based on a 10 °C Absolute Error (AE) threshold for the predicted temperature of the bearing.

A threshold in train speed was introduced in the rule just for scoring data sets reduction, and the expected subsequent accuracy of the rule's improvement. However, accuracy improvement was found not to be very significant for all cases.

To illustrate the difference in AE data distribution when the bearing is in good conditions versus when it comes to a degraded state, Figure 7 represents the temperature prediction AE distribution in periods of good (green) vs. degraded (blue) conditions, with train speeds  $TS_t \ge 90$  km/k.



Figure 7. Distribution of EA for good (green) and degraded condition (blue) periods, for a train speed  $TS_t \ge 90$  km/h. Taken from (Crespo Márquez, de la Fuente Carmona, et al., 2020).



### **Failure Mode Classification for CBM**

- The train axle bearing FM classification model is the second model contained in the DT of the CBM App in this case study.
- This modelling effort, to identify a certain bearing failure mode, required further ETL processes and different modeling tools. The most significant challenge was the decision (of the Smart Maintenance Department together with the Maintenance Engineering Department of the company), to approach this problem modeling temperature cycles instead of temperature points.
- This is a popular method (Healey et al., 2021) to study fatigue data analysis of mechanical components. In these cases, it is common to reduce a variable stress spectrum into a simpler, equivalent set of stresses. Methods that extract successively smaller cycles from a sequence are used to simplify the calculation of the fatigue life of a component from these simpler cycles (Healey et al., 2021).



### Failure Mode Classification for CBM

Determination of the following variables calculated from the extracted ones (Figure 8):

- Accumulated absolute error (Acc AE): This variable accumulates the AE when a positive is registered, since the first positive.
- Accumulated kilometers since the first positive: This is the total number of kms the train run since the first positive was registered.
- Accumulated kilometers in positive: This is the total number of the kilometers the train run in positive, since the first positive.



Figure 8. Sample data regarding Kms traveled in positive for different bearings



### **Failure Mode Classification for CBM**



Figure 9. Cycle count by varying the maximum distance between positives of the same cycle.



### **Failure Mode Classification for CBM**

Obtention of new following variables as per the cycle analysis performed:

- *Kilometers at the beginning of the cycle*: These are the kilometers that the bearing traveled, from the first positive, until a new cycle started.
- *Kilometers at the end of the cycle*: These are the kilometers that the bearing has traveled, from the first positive, until the end of the cycle.
- *Cycle Kms*: Kilometers that the train travels in a cycle (the cycle ends when the next positive is farther away from the previous one, than the limit in km established in each case).
- *Kilometers traveled between cycles*: These are the kilometers traveled between the end of one cycle and the beginning of the next one.
- *Cumulative cycles*: Cumulative number of cycles since the first positive.
- *Percentage of kilometers in active cycle*: Percentage of kilometers that the bearing accumulates in a cycle since the appearance of the first cycle.
- *Total kilometers in active cycle*: Accumulation of kilometers that the bearing run within cycles.
- Accumulated kilometers between cycles: This is the sum of the kilometers that a bearing traveled between cycles, up to the last cycle.
- Average of the kilometers between cycles: In this section we have the average of the kilometers traveled between cycles.
   Making this average gradually as we go from cycle to cycle.



### Failure Mode Classification for CBM



Figure 10. Sample of values obtained for cycle variables, when varying the maximum distance selected between positives of the same cycle.

**OMAINTE** 



### **Failure Mode Classification for CBM**

Although the main aim of the transformation process is to approximate the physical degradation model in a simpler way, it is observed that the amount of data to be considered and stored for the bearing analysis is also significantly reduced. The reduction achieved in the data to be stored per bearing studied is presented in Table 1.

Bearing Samples	REDUCTION of DATA POINTS for a				
	Maximum distance between positives of a cycle of				
	1km	5km	10km	20km	50km
KZ02 T3 AXLE 29	89.346	65	55	51	42
KZ15 T2 AXLE 1	78.318	416	281	207	152

Table 1. Reduction of the number of data records to be captured per bearing when applying the cycle algorithm.





### Failure Mode Classification for CBM

Table 2. Exam of an extrac with data fro several bearings, showing the number of da lines per bear (assuming 5 as max distar between positives of cycle)

	1	Bearing	Replaced	Damaged	Kms end of cycle	Kms between cycles	Cycle Kms	Cumulative Cycles	Kms in act cycle	%Kms in act cycle	Acc Kms betw Cycles	Avg Kms betw cycles
	2	KZ15 T0 EJE 17	1	1	345,6	2619,9	342,6	1,0	342,6	0,3%	2619,9	2619,9
	62	KZ15 T0 EJE 17	1	1	66704,7	58,5	31,7	61,0	10667,0	6,4%	56093,2	919,6
	63	KZ15 T1 EJE 22	1	1	25,1	8,0	22,0	1,0	22,0	0,0%	8,0	8,0
iple	201	KZ15 T1 EJE 22	1	1	92086,1	41,1	154,1	139,0	9694,0	5,0%	82430,0	593,0
1.	202	KZ15 T1 EJE 29	1	1	1211,8	13,4	1209,9	1,0	1209,9	1,2%	13,4	13,4
ct	459	KZ15 T1 EJE 29	1	1	105226,1	870,3	18,3	258,0	38379,0	18,7%	67715,5	262,5
	460	KZ13 T3 EJE 23	1	1	60,9	23,0	33,8	1,0	33,8	0,0%	23,0	23,0
m	743	KZ13 T3 EJE 23	1	1	110897,4	3865,3	0,8	284,0	12251,3	5,8%	102484,3	360,9
	744	KZ07 T3 EJE 1	1	1	79561,5	72,7	1,8	1,0	1,8	0,0%	72,7	72,7
	811	KZ07 T3 EJE 1	1	1	197818,7	130,8	3,3	58,0	2342,3	0,8%	116047,5	2000,8
	812	KZ02 T3 EJE 29	1	1	5,3	979,8	5,3	1,0	5,3	0,0%	979,8	979,8
	883	KZ02 T3 EJE 29	1	1	114651,1	2383,2	10,2	65,0	7770,8	3,6%	109263,4	1681,0
	884	KZ15 T2 EJE 1	0	0	82,7	8199,1	59,8	1,0	59,8	0,1%	8199,1	8199,1
e	1299	KZ15 T2 EJE 1	0	0	203641,3	7,7	3,6	416,0	27284,0	9,0%	176342,1	423,9
0	1300	KZ15 T2 EJE 2	0	0	4,2	18,6	3,4	1,0	3,4	0,0%	18,6	18,6
ata	1804	KZ15 T2 EJE 2	0	0	207253,8	508,7	0,0	505,0	32954,6	10,7%	174807,1	346,2
	1805	KZ15 T2 EJE 3	0	0	186233,4	2958,9	66,2	1,0	66,2	0,0%	189126,1	189126,1
ring	1949	KZ15 T2 EJE 3	0	0	245440,2	25,8	12,4	116,0	5735,5	2,3%	239730,5	2066,6
	1950	KZ15 T0 EJE 8	1	1	46,8	14,2	46,8	1,0	46,8	0,0%	14,2	14,2
km	1982	KZ15 T0 EJE 8	1	1	20361,6	7,9	4,7	33,0	2510,0	2,1%	17859,4	541,2
	1983	KZ15 T1 EJE 8	0	0	18,5	5,5	18,5	1,0	18,5	0,0%	5,5	5,5
nce	2059	KZ15 T1 EJE 8	0	0	184943,8	81,2	3,3	77,0	9383,9	3,3%	175641,1	2281,1
icc	2060	KZ15 T2 EJE 13	0	0	13,7	89,8	12,6	1,0	12,6	0,0%	89,8	89,8
	2937	KZ15 T2 EJE 13	0	0	270318,8	51,9	8,0	878,0	38209,0	10,3%	232160,5	264,4
	2938	KZ11 T1 EJE 15	1	0	9,8	10,1	9,8	1,0	9,8	0,0%	10,1	10,1
а	3072	KZ11 T1 EJE 15	1	0	123716,9	235,0	6,5	116,0	13641,2	6,1%	110310,6	951,0
a	3073	KZ15 T1 EJE 19	1	0	32,8	23543,6	32,8	1,0	32,8	0,0%	23543,6	23543,6
	3117	KZ15 T1 EJE 19	1	0	232583,5	44,1	6,3	36,0	1812,1	0,5%	230815,4	6411,5
	3118	KZ02 T2 EJE 27	1	0	41,6	46,0	41,5	1,0	41,5	0,0%	46,0	46,0
	3400	KZ02 T2 EJE 27	1	0	255001,7	6,4	197,9	248,0	13459,6	3,8%	241548,5	974,0
	3401	KZ02 T3 EJE 27	1	0	7,3	35,9	7,3	1,0	7,3	0,0%	35,9	35,9
	3463	KZ02 T3 EJE 27	1	0	319807,6	637,8	201,6	54,0	2890,3	0,7%	317554,9	5880,6
	3464	KZ16 T1 EJE 29	1	0	3,0	4619,3	3,0	1,0	3,0	0,0%	4619,3	4619,3
	3660	KZ16 T1 EJE 29	1	0	253375,3	8,0	1,6	171,0	7604,0	2,2%	245779,2	1437,3



### **Failure Mode Classification for CBM**

- Once the required data base is ready for model generation training and production, the process continues with the algorithm design, testing and validation.
- The algorithm attempts to separate bearings with internal deterioration from those with overtemperature caused by external causes, mainly due to the train axle guidance system problems. To that end, it is necessary to know the final diagnosis of all the bearings observed to have suffered overtemperature cycles. It is essential to have data on whether the bearing was replaced or not, and if once it was replaced, whether the analysis performed by the quality department found it with internal deterioration or not.
- Bearings in the train that were not replaced, but which had overtemperature cycles recorded, are obviously classified as "non-deteriorating" bearings. Basically, most of these bearings went back to normal temperature conditions when the train guidance problems were solved.
- The algorithm selected for this classification functionality can be chosen among different possibilities: according to its ROC curve (see Figure 11), classification error, gain, execution time, training time, etc. For this case, the selected algorithm has been Deep Learning.





### Failure Mode Classification for CBM

**Optimal Trade-offs between Complexity and Error** 



Used feature set Optimal trade-offs Original feature set Optimal feature set



Concerning final features selection for the model, notice that a complexity of 4 features achieved a lower error rate that the original selected set of 5 (that was also including the duration of the cycle, as feature).

So, the model is less complex and still more accurate than the original feature space (square in the graph). Using less features also means that models can be trained faster. The feature set which has been used to build the final model is shown.

21

Figure 12. Trade-offs between model dimensionality (complexity) and error, including final features selection and their weights (RapidMiner ®).





### Failure Mode Classification for CBM

Criterion	Value	STD
Accuracy	76.3%	± 0.2%
Classification error	23.7%	± 0.2%
AUC	91.6%	± 0.2%
Precision	100%	± 0.0%
Recall	5%	± 1.0%
F Measure	9.5%	± 1.9%
Sensitivity	5%	± 1.0%
Specificity	100%	± 0.0%

 Table 3. Classification algorithm performance metrics

	True range 1	True range 2	Class Precision
redicted range 1	785	248	75.99%
redicted range 2	0	13	100.00%
lass Recall	100.00%	4.98%	

**Table 4.** Sorting algorithm confusion matrix (range 1:Guidance FM; range2: Internal FM)



### **Data analytics for prognostics in CBM**

- Failure prognostics is defined (ISO 13381-1:2004) as "the Estimation of the Time to Failure (ETTF) and the risk of existence or later appearance of one or more failure modes". However, in most of the literature related to prognostics, the terminology Remaining Useful Life (RUL) is used, instead of ETTF (Medjaher, Tobon-Mejia, & Zerhouni, 2012).
- The concept of the RUL has been widely used in operational research, reliability, and statistics literature with important applications in other fields such as material science, biostatistics, and econometrics. Clearly the definition of the useful life depends on the context and operational characteristics (Si, Wang, Hu, & Zhou, 2011).
- Concerning the estimation of the RUL, the existing approaches fall into three main categories (Jardine, Lin, & Banjevic, 2006): statistical approaches, artificial intelligence (AI) approaches and model-based approaches.



### **Data analytics for prognostics in CBM**

- In this case study a statistical approach is followed to estimate the RUL (of any bearing of a train), once a positive (or anomaly detected for a failure mode) appears in a train axle bearing.
- A positive (according to the Procedure for the Design and Implementation of CBM Plans in the company) is defined as the occurrence of an absolute error (AE) of prediction greater than 10°C between the actual bearing temperature and that predicted by the ANN designed for detection, when the train is running at more than 90 km/h (i.e., AE ≥ 10°C, TS ≥ 90 km/h) and for more than one minute.
- RUL is now defined as a random variable that, estimated from the appearance of the first positive, offers a good prediction of the life of the element until its replacement due to over temperature or noise. This replacement was performed after the activation of the safety alert in the train monitoring and control system (TCMS) and/or because of a certain inspection (probably during a weekly train inspection in the workshop). The safety alert is triggered when the temperature difference between the four bearings of the same axle is higher than  $25^{\circ}C (Tmax Tmin) \ge 25^{\circ}C and$  this condition is maintained for more than 1 minute.
- Company's objective through the analysis included in this part of the case study is to foresee the recommended time of bearing replacement, after its first positive, even without prior inspection, according to statistical estimates.



### **Data analytics for prognostics in CBM**

To calculate the RUL at point A, is necessary to model the random variable "PF interval", i.e. the interval (in time, km, or representative unit of measurement) that elapses between the first positive (point P, agreed in the CBM procedure for:  $AE \ge$  10.*C*, TS  $\ge$  90 *km/h*) and the possible replacement due to overtemperature and/or noise of a bearing.

The point F considered takes place, in general, after the activation of the safety alert in the train monitoring and control system (TCMS), this condition is not of functional loss of the bearing, but of operation in conditions of lower safety level. Then it is possible to define, for this case study:

RUL = RUL AF interval = (PF Interval - PA Interval).

The determination of the RUL will be made from the estimation of the distribution function of the PF interval, using a statistical technique such as the Weibull analysis.



Figure 13. P-F Curve and P-F time interval. Estimated time to failure (RUL)



### Interaction with the CBM DT

- The functionality of the DT allows the evaluation of the failure mode risk level and the subsequent control actions, this will allow the maintenance staff to schedule convenient maintenance activities.
- Interaction with the DT must be done using simple business rules resulting in a practical business process.
- Any new event detected by the DT leading to a new state of the asset concerning a failure mode will be a call for maintenance action.
- For the correct interpretation of the method of interaction with the DT, Table 6 describes the necessary concepts to be reviewed (taken from an original work in Martínez-Galán Fernández, Guillen López, Crespo Márquez, Gómez Fernández, & Marcos, 2022).



### **Interaction with the CBM DT**

#### Concept

#### Event

Recordable, scheduled, or supervening time, at which the risk level of the affected failure modes must be reanalyzed.

#### State

Qualitative level of risk at a given time. Each event causes a possible change in the level of risk.

#### **Failure Mode**

Failure modes involved that can be fully or partially managed by CBM. Monitoring solutions and maintenance tasks are applied at failure mode level.

Types	
-------	--

- Monitoring Event: Events taking place because of the CBM App (and its DT algorithms).
   They can be detection events, diagnostics events, or prognostics events.
- Preventive Maintenance Events: Maintenance programmed or unforeseen events. They can be for example inspections or any PM activity.
- Fault: State after the failure has occurred. State in immediate replacement or repair of the item is required.
- High Risk: State of operation closest to failure. Short-term activities are scheduled to reduce the level of risk.
- Medium Risk: State in which an anomaly has been detected but with some security it is possible to continue operating under normal conditions. Medium-term activities are planned to confirm the risk and analyze how it evolves.
- Low Risk: Normal operating state of the item
- Primary failure mode (PFM)\*
- Secondary failure mode (SFM)\*: initiated by a PFM

\* Terminology adopted from ISO 13381, (ISO, 2015)

**Table 6.** Key concepts in DT interaction with maintenance techs. Adapted from (Martínez-GalánFernández, Guillén López, Márquez, Gómez Fernández, & Marcos, 2022)





### **Interaction with the CBM DT**

To describe these concepts in a graphical manner, a CBM sequence affecting two failure modes is pictured in Fig. 14. In this case, Monitoring Events and PM Events may change the each one of the FMs risk level (FM1: Internal degradation and FM2: External guidance failure).



**Figure 14.** Graphic representation of the CBM APP DT interaction with Maintenance technicians. Adapted from (Martínez-Galán et al., 2022).





### **Fulfillment of DT REQUIREMENTS**

Requirement	Case study description
Scalability.	The DT model has been scalable to all train bearings requiring only the development of models per axle bearing position, regardless the axle in the train nor the train in the fleet.
Interoperability	Data used to train the three different types of models came from the same source and there is a procedure explaining how original data is converted and matches the different predictive analytics data models. Real time data is now used to generate an on-line output;
Expansibility.	There is a clear possibility to integrate new models. For instance, RUL models based on machine learning models have been introduced to replace statistical models in some applications with more consistent data.
Fidelity	The ML models for anomalies detection replace in this case, with high tested precision, the very complex physical models related to the calculation of the dynamic behavior of loads in the train per axle bearings in each railway point at a certain speed.

**Table 7.** CBM App DT Fulfillment of the six requirementsextracted from those found in the DT literature.





Interaction

**THE 20<sup>TH</sup> INTERNATIONAL OPERATIONS & MAINTENANCE CONFERENCE IN THE ARAB COUNTRIES** 

### **Fulfillment of DT REQUIREMENTS**

Requirement Case study description This part has been found a very interesting requirement to fulfill. When modeling a given failure mode (FM) different risk levels or states are proposed: low, medium, high and fault. At the same time two different types of events may show up: monitoring and preventive maintenance events. It is considered that both monitoring events and PM events (with human intervention) may lead to a change in the risk level of one or more failure modes of the asset. This is because these events trigger a new risk assessment of the affected FMs. A given event may affect different failure modes and in different ways. It is also assumed that reaching a new failure mode state triggers a maintenance action (the release of an algorithm for detection or prediction, an inspection, a replacement, etc.). This human supervision of the model's performance and interaction with the DT resulted to be critical for the DT success. Integration The DT is to be integrated in the App in place, to control the trains fleet dynamic maintenance. Axle

bearing DT must be incorporated into the comprehensive train CBM App. In this App, a total of 10 train critical systems are monitored to generate an on-line train risk assessment and to suggest an immediate action. Understanding the implications of each system risk, according to each system criticality, is critical to establish an effective dynamic maintenance strategy. In this case this DT has been integrated within Google cloud infrastructure/services.

> **Table 7.** CBM App DT Fulfillment of the six requirements
>  extracted from those found in the DT literature (cont.).

/OMAINTECConf



### CONCLUSIONS

- We have seen how to use the DMM framework for the functional definition of a DT designed to support a CBM application based on predictive analytics.
- We presented the data models for each predictive analytics algorithm.
- We described the information that the end user exchanges with the App and how this interaction takes place. Moreover, we verified that the design of the DT also meets other requirements: scalability, interoperability, expandability, fidelity in integration with existing dynamic maintenance management tools have also been contrasted.
- It is proposed that tools of this type (DTs) should be documented using a scheme like the DMM, and controlled according to the presented DT requirements.
- In fact, using this framework, any tool used in intelligent applications for maintenance management can be defined, not only those with an important operational nature such as the CBM, but also others of a more strategic nature such as those for criticality analysis or those for long-term asset health analysis.



# THANK YOU!

#### **Prof. Adolfo Crespo Marquez** Full Professor, University of Seville, Spain

**GOGO** #OmaintecConf

An Initiative by



International Grou جموعة أكزيكون الدولية International Business Un

Organized by