

Modelización híbrida para el diagnóstico y pronóstico de fallos en el sector del transporte. Datos adquiridos y datos sintéticos

■■■■
Madhav Mishra, Urko Leturiondo-Zubizarreta,
Óscar Salgado-Picón y Diego Galar-Pascual
Luleå University of Technology y IK4-Ikerlan

DOI: <http://dx.doi.org/10.6036/7252>

1. INTRODUCCIÓN

El mantenimiento basado en la condición (CBM, del inglés *condition based maintenance*) es un tipo de mantenimiento para equipos cuyo uso ha crecido exponencialmente en los últimos años, en gran medida por su capacidad para predecir fallos futuros, aumentando la fiabilidad y la seguridad del activo monitorizado. El CBM cuenta con dos procesos fundamentales: la diagnosis y la prognosis [1,2]. La diagnosis pretende detectar los fallos que se producen a través de la aparición de los correspondientes modos de fallo y, una vez han sido detectados, éstos deberán ser identificados y vigilados de cara a realizar un análisis de la salud real del sistema. Por su parte, se llama prognosis al acto de realizar un pronóstico. Este concepto, aplicado al campo del mantenimiento, radica en determinar el tiempo residual de vida útil (RUL, del inglés *remaining useful life*) de un equipo o componente [3]. La diagnosis y la prognosis constituyen el binomio del monitorizado de la condición o *condition monitoring*, donde la optimización del proceso de diagnosis es clave para obtener un buen resultado del proceso de prognosis [4].

En los sistemas del sector del transporte, donde la interacción de cientos de componentes y subsistemas está involucrada en la operación, resulta difícil disponer de un completo catálogo de fallos correspondiente a todos los sistemas

que operan de manera individual y a las distintas composiciones cuando éstos actúan de manera conjunta y producen nuevos fallos. Ello conlleva que gran parte de los fallos pueden no ser detectables y los que lo son pueden no ser aislables o convenientemente vigilados, como se puede observar en la Fig. (1).

Por otro lado, aquellos modos de fallo que pueden ser vigilados a través de indicadores de condición que muestran los síntomas asociados son controlados

y su frecuencia de fallo se mitiga con diseños robustos y/o mantenimientos conservadores. En el sector del transporte el coste se convierte en una variable secundaria, al igual que sucede en otros sectores tales como el de la energía o el de la química, donde la seguridad prima por encima de todos los factores de eficacia y eficiencia para evitar cualquier tipo de accidente [5,6]. Esto obliga a reemplazos prematuros de componentes obligados por regulaciones estrictas. De hecho, la regulación existente en el sector del transporte, aplicada primero en el sector aeronáutico y ahora trasladada al ferroviario, obliga a aplicar criterios de fiabilidad, disponibilidad, mantenibilidad y seguridad (RAMS, del inglés *reliability, availability, maintainability and safety*) muy restrictivos, tal y como marca, por ejemplo, la norma EN 50126-5.

Este cambio paradigmático en el entorno de la prognosis de sistemas de transporte exige nuevos enfoques ya

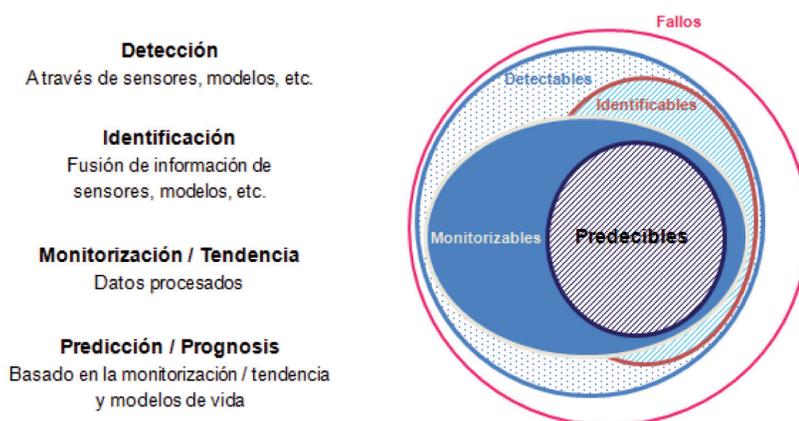


Fig. 1: Procesos de detección y predicción de fallos

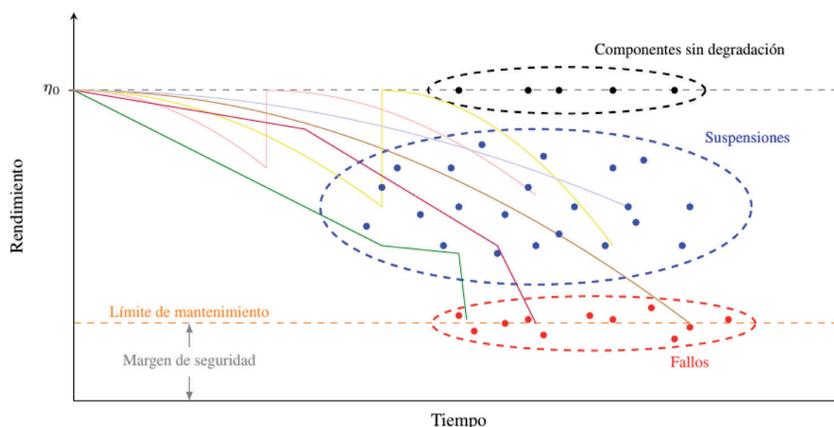


Fig. 2: Evolución del rendimiento de un sistema a lo largo del tiempo

que la realización de pronósticos basados en patrones de degradación [7], como muestra la Fig. (2), donde los modos de fallos y sus evoluciones eran parcialmente conocidas a priori, deja de tener sentido.

El escenario real al que el analista se enfrenta es aquél donde sólo un número de componentes o subsistemas fallan debido a muertes prematuras o a elementos que deliberadamente se dejan operar hasta que se produce el fallo debido a su baja criticidad [8]. Sin embargo, todos los elementos críticos que pueden afectar a la funcionalidad y, especialmente, a la seguridad del sistema se reemplazan en épocas prematuras de su degradación (suspensiones). Esto conlleva la no disponibilidad de patrones completos de dicha degradación, ni siquiera cercanos al límite de mantenimiento, establecido según distintas normativas, como es la EN 13848-5 para el caso ferroviario. Esta falta de datos viene también dada porque unidos a los reemplazamientos prematuros a veces se realizan reemplazos oportunistas.

Por todo ello, en este artículo se propone que la prognosis de sistemas de transporte debe integrar retazos de información de diversas fuentes considerando que una base histórica de fallos es inviable por la sobreprotección y sobremantenimiento del sistema en un entorno fuertemente regulado que no permite deterioros y asegura fiabilidades muy altas. Estos retazos de información son las degradaciones parciales de los sistemas hasta su reemplazo, el conocimiento físico de los mismos y por supuesto la integración en el proceso de fallos que pudieran producirse y que no se tuvieron en cuenta en etapas de diseño. Todo ello debe contribuir a la estimación de una RUL donde la falta de información histórica es un hándicap y la información de la fase de diseño con la información de campo deben ser fusionados de tal manera que la predicción arroje unos intervalos de confianza lo suficientemente buenos como para tomar decisiones de mantenimiento y operación en base a ellos.

2. MODELIZACIÓN CLÁSICA DE SISTEMAS

Para llevar a cabo el CBM se pueden emplear modelos físicos, modelos basados en datos o modelos simbólicos.

Los modelos físicos se construyen desde el conocimiento de la física del sistema a modelizar y vienen expresados a través de ecuaciones [9]. Estas ecuaciones, que pueden ser tanto ecuaciones diferenciales ordinarias o ecuaciones en derivadas parciales, pueden clasificarse de la siguiente manera:

- Ecuaciones de balance (por ejemplo, reacciones químicas).
- Ecuaciones de estado físicas o químicas (por ejemplo, ecuaciones que relacionan variables de estado de un sistema).
- Ecuaciones fenomenológicas (por ejemplo, la ley de *Fourier* de conducción térmica).
- Ecuaciones de interconexión (por ejemplo, la ley de corrientes de *Kirchhoff*).

Una vez que se definen las ecuaciones, el modelo físico está construido. Dichas ecuaciones pueden ser simplificadas a través de linealizaciones, aproximaciones y reducciones de orden, entre otras técnicas [9], de manera que se consigue que la resolución matemática sea más factible.

Los modelos físicos son muy útiles para describir el comportamiento de sistemas variantes en el tiempo, teniendo en cuenta diferentes modos de operación, transitorios y condiciones medioambientales variables. Además, la modelización de los distintos modos de fallo es clave a la hora de obtener una correcta representación del comportamiento de los sistemas [10,11]. Sin embargo, a medida que se incluyen efectos y la complejidad de un modelo crece, el esfuerzo precisado para desarrollarlo y validarlo es mayor. Esto conlleva una necesidad mayor en lo que a recursos computacionales hace referencia. Es por ello por lo que se requiere definir un límite de complejidad en el modelo.

La principal ventaja de los modelos físicos respecto a los modelos basados en datos consiste en el anteriormente mencionado conocimiento de la física del sistema monitorizado [12]. Los modelos basados en datos pierden la relación causal entre datos y el mundo físico, poniendo en cuestión la fiabilidad de dichos algoritmos. Varias técnicas para la modelización basada en datos están descritas en [1,2].

Por último, los modelos simbólicos [13] emplean especificaciones técnicas

de los componentes, recomendaciones y tests de los fabricantes, órdenes de trabajo, informes de mantenimiento y experiencia aportada por los responsables de mantenimiento. Esta información puede encontrarse en forma numérica y, más comúnmente, en forma de palabras. En este último caso cabe discernir dos situaciones: si se trata de información escrita se puede emplear para crear relaciones causales de desarrollo de fallos; sin embargo, si la información se transmite verbalmente es más complicado obtener descripciones detalladas de dependencias complejas y de comportamientos variantes en el tiempo.

3. MODELIZACIÓN HÍBRIDA

La modelización híbrida que se propone en este artículo emplea la fusión de datos provenientes de diferentes fuentes con el objetivo de llevar a cabo un análisis de prognosis de un sistema. Este tipo de modelización cuenta con algunas ventajas, como puede ser en el caso de una aeronave. Se pueden extraer datos de una aeronave cuando el activo está operando sin ningún tipo de fallo. Sin embargo, cuando un indicador clave de rendimiento (KPI, del inglés *key performance indicator*) del sistema sobrepasa el umbral de mantenimiento preestablecido se llevan a cabo las tareas de mantenimiento pertinentes. Esto implica que únicamente se pueden adquirir datos hasta un punto cercano al instante t_m , que representa el instante de tiempo en el que dicho límite es sobrepasado, como se muestra en la Fig. (3). Tras el umbral de mantenimiento se encuentra el umbral de servicio, que corresponde al límite en el cual el sistema no se encuentra en condiciones aptas para su funcionamiento. Es por ello por lo que se pretende evitar que el sistema llegue a este límite (que se alcanzaría en un instante de tiempo t_{ser}). Finalmente, no se permitirá en ningún momento que el indicador exceda el umbral de seguridad predefinido (que se alcanzaría en un instante de tiempo t_{seg}) por las siguientes razones:

- Seguridad: algunos fallos pueden poner en riesgo tanto al sistema como a humanos.
- Coste: el desarrollo de un fallo en un componente de una aeronave puede ser muy caro.

- Cuestiones medioambientales: el efecto de un fallo puede ser perjudicial para el medio ambiente.

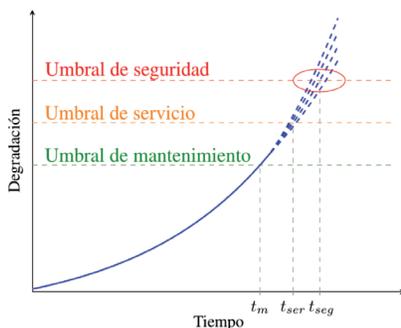


Fig. 3: Análisis de la tendencia para la estimación de la vida útil remanente

En consecuencia, no se pueden adquirir datos del sistema real bajo ciertas condiciones de fallo, de forma que sólo se pueden obtener datos del estado sano (S) del sistema, como se muestra en la Fig. (4) teniendo en cuenta dos indicadores. Sin embargo, las distintas condiciones de fallo que pueden producirse en el sistema pueden ser simuladas empleando modelos físicos. Para ello, los modos de fallo correspondientes al sistema han de ser previamente definidos empleando diversas técnicas como el mantenimiento centrado en la fiabilidad o el análisis modal de fallos y efectos, entre otras. Una vez que se han determinado los modos de fallo y se han modelizado en el modelo físico, se pueden generar datos mediante simulaciones, obteniendo “datos sintéticos”. En la Fig. (5) se muestran datos correspondientes al caso sano (S) y a varios casos de fallo (F_1 , F_2 y F_3) obtenidos mediante modelos físicos.

Así, el conjunto de datos final contará con datos adquiridos del sistema real cuando éste se encuentra en condiciones sanas y datos sintéticos generados

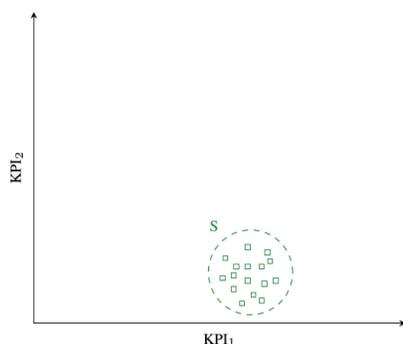


Fig. 4: Datos sanos adquiridos del sistema real

con un modelo físico que caracteriza los distintos modos de fallo identificados en el sistema. A la combinación de un modelo basado en datos y un modelo físico se le denomina modelización híbrida. Una vez que se cuenta con los datos obtenidos del sistema real y de las simulaciones realizadas mediante el modelo físico, se procede a clasificar dichos datos y a realizar un ajuste para optimizar dicha clasificación.

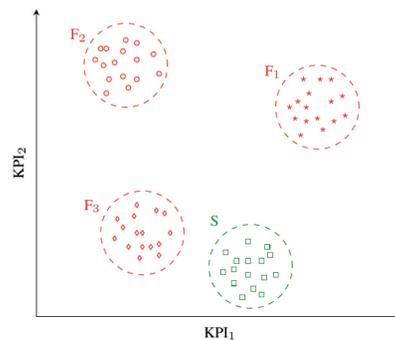


Fig. 5: Datos sanos adquiridos del sistema real y datos sintéticos generados mediante el modelo físico

3.1. PROCESO DE CLASIFICACIÓN

En esta metodología se va a emplear la clasificación supervisada, también conocida como aprendizaje automático, la cual parte de una situación en la que se cuenta con datos donde se conoce a qué grupo pertenece cada individuo y cuyo objetivo es encontrar una relación entre los individuos y los grupos de manera que sea posible asignar un grupo a un nuevo individuo.

Esta elección se realiza debido a que se cuenta con la información de la tipología de los grupos a los que pertenecen los datos. Además, se aprovecha que las distintas técnicas de este tipo de clasificación (máquinas de vector de soporte, el método de los k vecinos más próximos o redes neuronales, entre otras) no son capaces de discernir entre las distintas fuentes de datos de entrada, de forma que se trata de una clasificación “ciega”.

3.2. PROCESO DE AJUSTE

Una vez que el proceso de clasificación se ha llevado a cabo, los datos nuevos que se adquieran del sistema pueden ser fácilmente categorizados a alguno de los grupos definidos. Sin embargo, puede darse el caso en el que

haya datos que no se corresponda con ninguno de dichos grupos. A este estado se le denomina fallo desconocido (NFF, del inglés *non fault found*). La Fig. (6) muestra una gráfica donde se produce esta situación. Los datos que no corresponden a ninguno de los grupos predefinidos (S , F_1 , F_2 y F_3) son considerados NFF. Existen dos razones principales por las que pueden aparecer datos de tipo NFF:

- El modelo físico no es sensible a uno de los modos de fallo estudiados, de forma que los datos adquiridos no se asemejan a los datos generados mediante el modelo físico.
- Los datos de tipo NFF no corresponden con ninguno de los modos de fallo identificados.

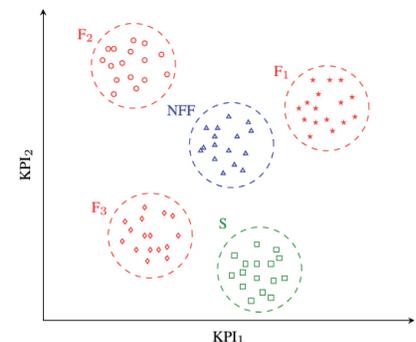


Fig. 6: Aparición de datos del tipo NFF

La aparición de este tipo de datos ha de aprovecharse para actualizar la clasificación realizada y, así, completarla. Los datos de tipo NFF se consideran nuevas categorías relacionadas con otros modos de fallo y, junto a los datos

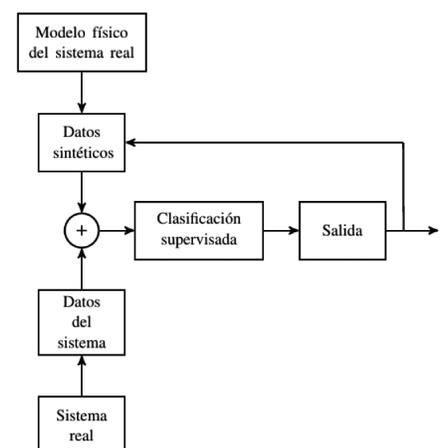


Fig. 7: Proceso de ajuste del sistema de clasificación

obtenidos del sistema y los datos sintéticos generados, se repite la clasificación supervisada. Al proceso de actualización automática del criterio de clasificación se le llama proceso de ajuste. Una vez que el proceso de ajuste es llevado a cabo se obtiene un mejor conocimiento de la evolución de los fallos, realizando así de forma más sencilla el proceso de pronóstico del sistema. La Fig. (7) muestra un esquema de este proceso.

4. DE FUSIÓN DE DATOS HACIA FUSIÓN DE INFORMACIÓN EN EL SECTOR FERROVIARIO

La metodología propuesta en el apartado anterior se basa en la fusión de datos, los cuales provienen de dos fuentes: del propio sistema monitorizado y de la modelización física del mismo. Sin embargo, a estos datos hay que añadirles modelos simbólicos, los cuales aportan información acerca de los activos y de su gestión. Dicha información procede de órdenes de trabajo o de diversas fuentes de mantenimiento y/o gestión, como son el *software* de supervisión, control y adquisición de datos (SCADA, del inglés *Supervisory Control And Data Acquisition*), los sistemas de planificación de recursos empresariales (ERP, del inglés *Enterprise Resource Planning*) o la gestión de mantenimiento asistido por computadora (CMMS, del inglés *Computer Maintenance Management Software*), entre otros [14]. La información procedente de dichas fuentes está basada en la experiencia de los operarios y de los responsables de mantenimiento y gestión que controlan el estado de los activos, así como de los fabricantes de los componentes o sistemas.

El sector del transporte ha sido especialmente sensible a todos los avances en diagnóstico y pronóstico debido a que los usuarios demandan calidad en servicio sin comprometer la seguridad de los sistemas. Es también cierto que los sistemas de transporte han incrementado su complejidad con elementos mecatrónicos y la intrusión del *pervasive computing* en el sector. Todo ello ha generado una ingente cantidad de información que puede ser empleada para una mejor estimación de la vida útil de sistemas, subsistemas y componentes.

Este tipo de información es abundante en el campo ferroviario, tal y como se muestra en la Fig. (8), donde existen numerosos sistemas, como el propio material rodante hasta la vía y los elementos de señalización.

Por un lado, existen modelos físicos para la degradación de componentes o subsistemas como carril, rueda, traviesas e incluso catenarias o pantógrafos. Por otro lado, diversas técnicas de monitorización de la condición embarcadas o en vía proveen datos sobre la geometría del carril, de las ruedas, ejes, calidad del balasto, etc., unido con ensayos no destructivos realizados regularmente. A estos datos se une la información referente a informes de mantenimiento de cada subsistema, la información de stock de reemplazos, las planificaciones de mantenimiento existentes, la hoja de ruta del maquinista, el horario de via-

jes, las órdenes de trabajo, las leyes y directrices del ámbito ferroviario, las normas, etc. Estas últimas forman un conjunto de información que, aunque se trate de una naturaleza diferente a la de los datos adquiridos o sintéticos, sirve para complementar al resto de datos.

Al empleo conjunto de datos (adquiridos o sintéticos) y modelos simbólicos se le denomina fusión de información. En general, la fusión de datos y de información, la cual se muestra gráficamente en la Fig. (9), permite a los responsables de mantenimiento del tren y de la infraestructura ferroviaria en la que circula tener un mejor conocimiento del estado de los diferentes activos y poder gestionar y establecer las necesidades de mantenimiento existentes para que dichos activos operen sin fallos y se mantenga la seguridad del tren así como la de sus pasajeros.

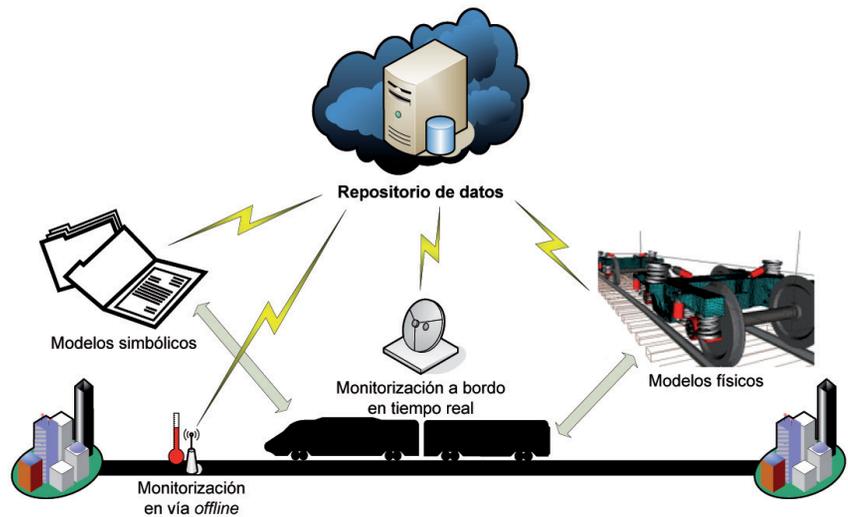


Fig. 8: Origen de datos en el sector ferroviario

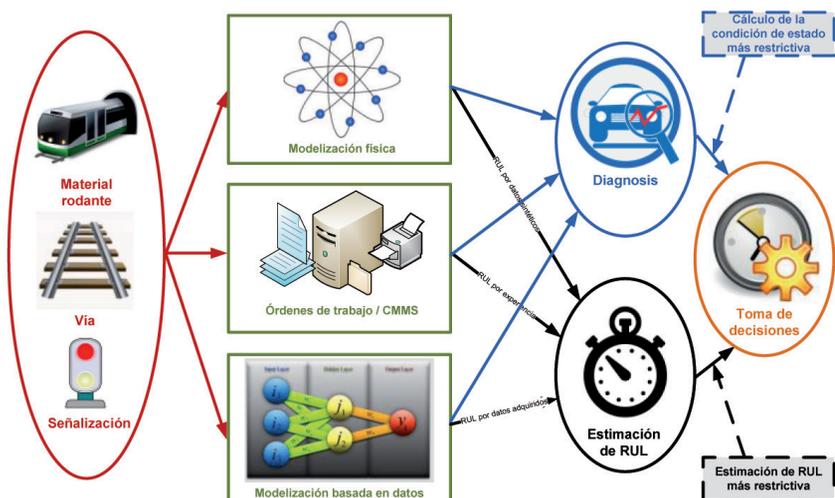


Fig. 9: Fusión de información para la toma de decisiones [14]

5. CONCLUSIONES

En este artículo se ha presentado una metodología para mejorar los procesos de diagnóstico y pronóstico. La carencia de datos correspondientes a fallos en operación o accidentes en algunos sistemas supone una circunstancia desfavorable a la hora de predecirlos. Es por ello por lo que la generación de datos sintéticos a través de modelos físicos que cubran un abanico de modos de fallo y su acoplamiento con los datos existentes de los sistemas aporta un conocimiento amplio de las posibles condiciones del sistema objeto de estudio. Por otra parte, en el caso en el que el sistema se comporte de una forma inesperada, ya sea porque se esté produciendo un modo de fallo no contemplado o porque la modelización no es sensible a alguno de los modos de fallo, se plantea que los datos correspondientes a dicho estado se utilicen para reentrenar y mejorar el sistema de clasificación.

Se ha visto, además, que esta metodología tiene un especial interés en sistemas complejos, ya que es más difícil predecir los fallos en sistemas de estas características. Así, esta metodología supone un paso adelante en el conocimiento del estado de sistemas donde la seguridad es crítica, como es el caso del sector transporte en general y del sector ferroviario en particular.

AGRADECIMIENTOS

Este artículo ha sido financiado en parte por el proyecto ETORTEK MECOFF, financiado económicamente por el Gobierno Vasco bajo el proyecto No. IE13-379. Los hallazgos, interpretaciones y conclusiones expresados en este documento son de los autores y no reflejan necesariamente las posturas de las entidades que lo financian. Los autores también quieren dar las gracias a LTU-SKF University Technology Centre (UTC) y a la Fundación de Centros Tecnológicos - Iñaki Goenaga.

PARA SABER MÁS

- 7 p.1483-1510. DOI: <http://dx.doi.org/10.1016/j.ymssp.2005.09.012>
- [2] Heng A, Zhang S, Tan A C C et al. "Rotating machinery prognostics: State of the art, challenges and opportunities". *Mechanical Systems and Signal Processing*. Vol.23-3 p.724-739. DOI: <http://dx.doi.org/10.1016/j.ymssp.2008.06.009>
- [3] Dragomir O E, Gouriveau R, Dragomir F et al. "Review of prognostic problem in condition-based maintenance". En: *Proceedings of European Control Conference, ECC09 (Budapest, 23-26 de agosto de 2009)*. 2009.
- [4] Xue F, Bonissone P, Varma A et al. "An instance-based method for remaining useful life estimation for aircraft engines". *Journal of Failure Analysis and Prevention*. Vol.8-2 p.199-206. DOI: <http://dx.doi.org/10.1007/s11668-008-9118-9>
- [5] Evans A W. "Fatal train accidents on Europe's railways: 1980-2009". *Accident Analysis & Prevention*. Vol.43-1 p.391-401. DOI: <http://dx.doi.org/10.1016/j.aap.2010.09.009>
- [6] Ayres J, Shirazi H, Carvalho R et al. "Modelling the location and consequences of aircraft accidents". *Safety Science*. Vol.51-1 p.178-186. DOI: <http://dx.doi.org/10.1016/j.ssci.2012.05.012>
- [7] Zoeteman A, van Meer G. "A yardstick for condition-based and differentiated planning of track and turnout renewal: a major step towards full decision support". En: *Proceedings 7th World Congress on Railway Research (Montréal, 4-8 de junio de 2006)*. 2006. P.1-11.
- [8] Santos J, Wysk R, Torres J. *Mejorando la producción con lean thinking*. 1ª edición. Madrid: Pirámide, 2010. 296p. ISBN: 978-84-368-2422-3
- [9] Isermann R, Münchhof M. *Identification of dynamic systems. An introduction with applications*. 1ª edición. Berlin: Springer-Verlag, 2011. 705p. ISBN: 978-3-540-78878-2
- [10] Li Y, Kurfess TR, Liang SY. "Stochastic prognostics for rolling element bearings". *Mechanical Systems and Signal Processing*. Vol.14-5 p.747-762. DOI: <http://dx.doi.org/10.1006/mssp.2000.1301>
- [11] Chen Z, Shao Y. "Dynamic simulation of spur gear with tooth root crack propagating along tooth width and crack depth". *Engineering Failure Analysis*. Vol.18-8 p.2149-2164. DOI: <http://dx.doi.org/10.1016/j.engfailanal.2011.07.006>
- [12] Luo J, Namburu M, Pattipati K et al. "Model-based prognostic techniques". En: *Proceedings AUTOTESTCON 2003. IEEE Systems Readiness Technology Conference (Anaheim, 22-25 de septiembre de 2003)*. 2003. P.330-340. ISBN: 0-7803-7837-7
- [13] Galar D, Kumar U, Villarejo R et al. "Hybrid prognosis for railway health assessment: an information fusion approach for PHM deployment". *Chemical Engineering Transactions*. Vol.33 p.769-774. DOI: <http://dx.doi.org/10.3303/CET1333129>
- [14] Galar D, Kumar U. "SMART bearings: From sensing to actuation". En: *Proceedings of the 12th IMEKO TC10 Workshop on Technical Diagnostics: New Perspectives in Measurements, Tools and Techniques for Industrial Applications (Florenca, 6-7 de junio de 2013)*. 2013. P.21-30. ISBN: 978-88-903149-8-8