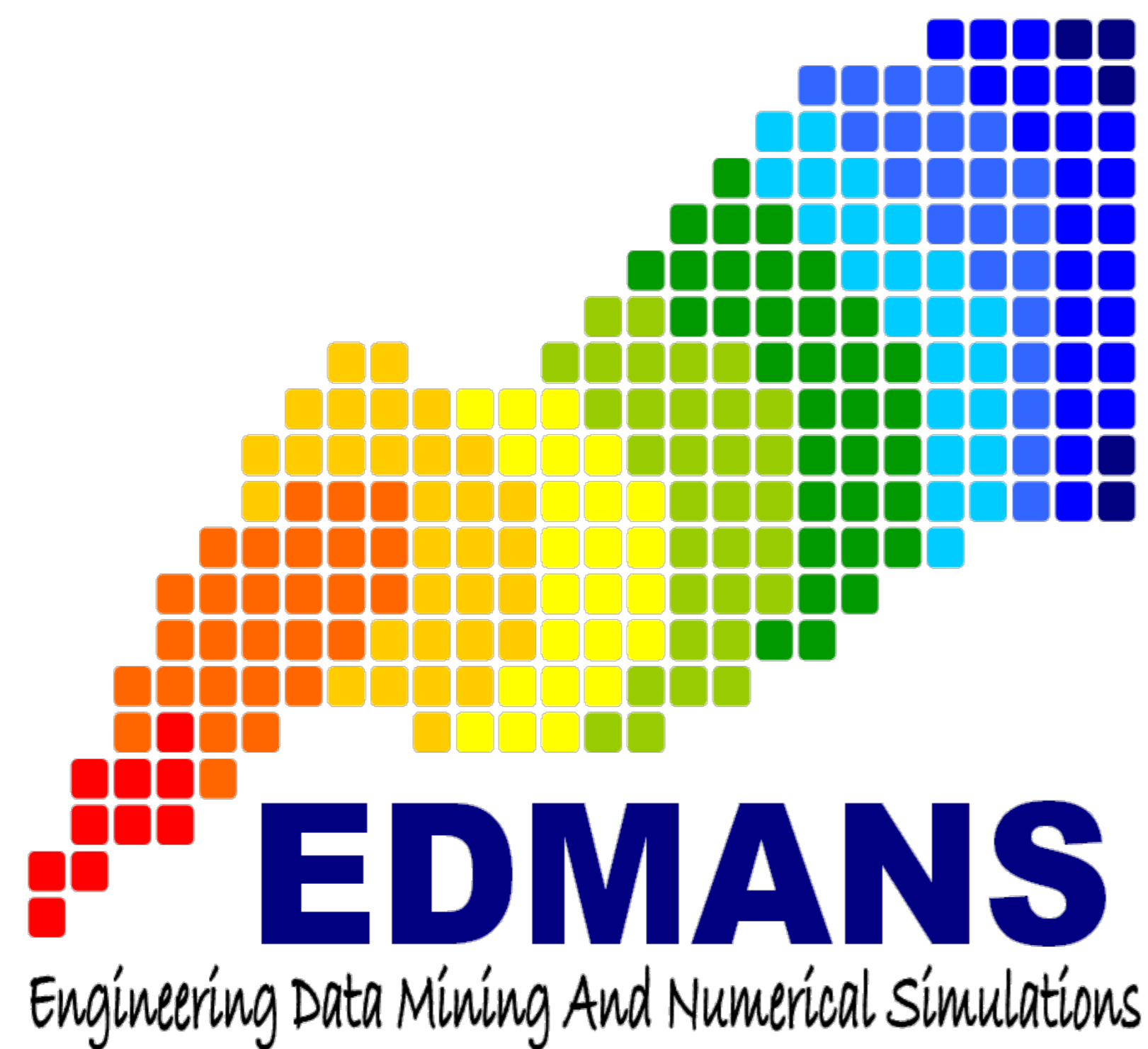


# Industrial Big Data (IBD)

Prospectiva aportaciones Tecnológicas:  
Píldoras de Estado de la Técnica hacia la Industria 4.0



**UNIVERSIDAD  
DE LA RIOJA**

**Fco. Javier Mtnez de Pisón Ascacibar**

**EDMANS Group**

**fjmartin@unirioja.es**

***<http://mineriadatos.com>***

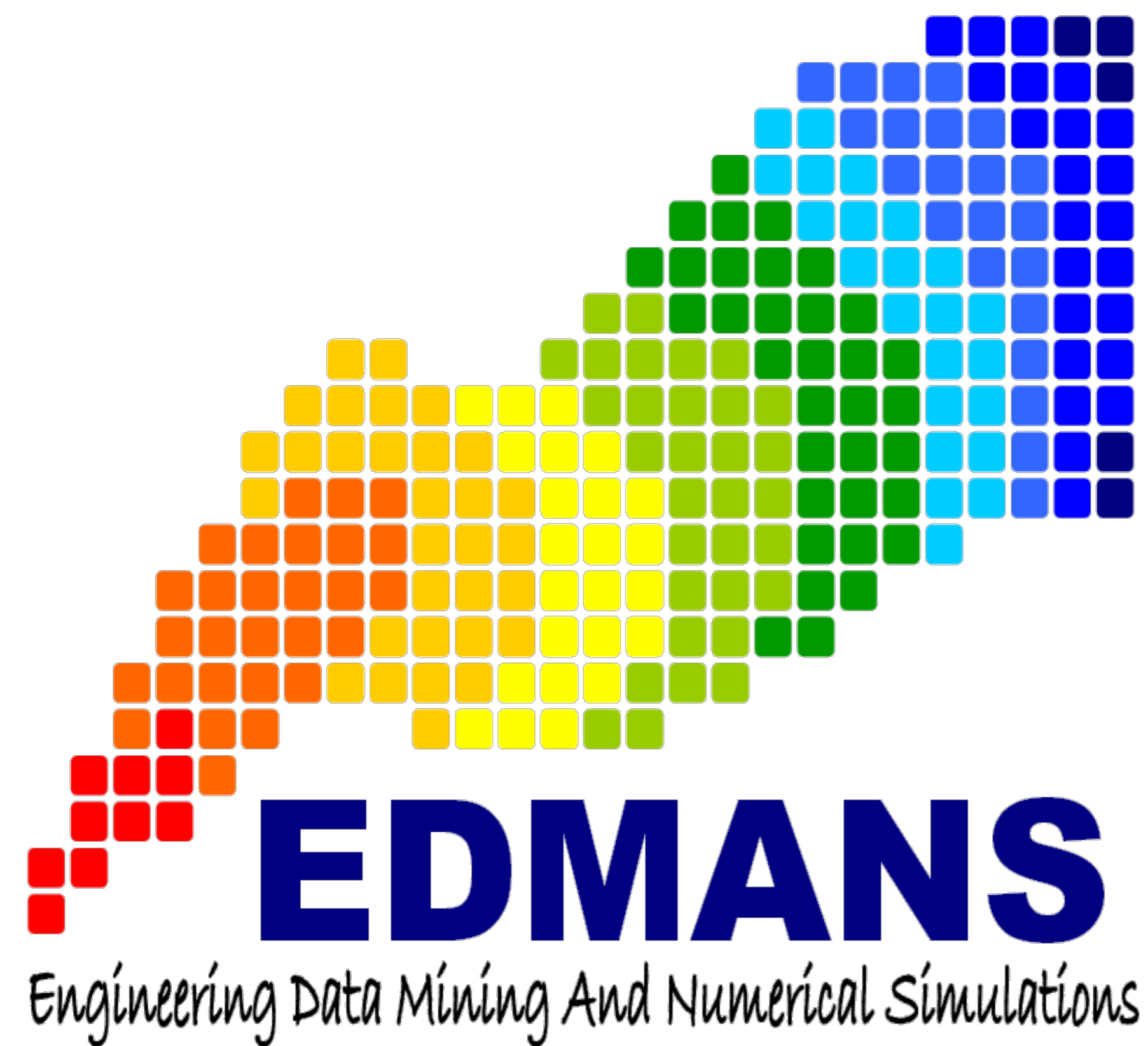


INDUSTRIA CONECTADA 4.0



**Gobierno  
de La Rioja**

[www.larioja.org](http://www.larioja.org)



Desde 1997, el grupo de investigación EDMANS (*Engineering Data Mining And Numerical Simulations*) está formado por ingenieros especializados en diferentes campos de la ingeniería (informática, calidad, estadística, automatización, modelizado, diseño de máquinas, procesos de fabricación, diseño de estructuras, medio ambiente, etc.). Este grupo, desarrolla su actividad de investigación en diversos proyectos europeos, nacionales y con empresas, fundamentalmente en estos dos ámbitos:

- 1. El análisis y mejora de procesos industriales, empresariales y medioambientales mediante técnicas de Big Data.**
- 2. El estudio mediante elementos finitos (FEM) y dinámica de fluidos computacional (CFD) de procesos mecánicos, térmicos o de fluidos en sistemas mecánicos, estructurales o procesos medioambientales.**

El objetivo fundamental es llegar a obtener  
una **mejor comprensión de los procesos estudiados**  
con el objetivo de **mejorar la calidad** de los mismos y los productos que generan.

[Home](#)[News](#)[Members](#)[Publications](#)[Projects](#)[Thesis](#)[Patents](#)

## EDMANS Research Group

Engineering Data Mining And Numerical Simulation



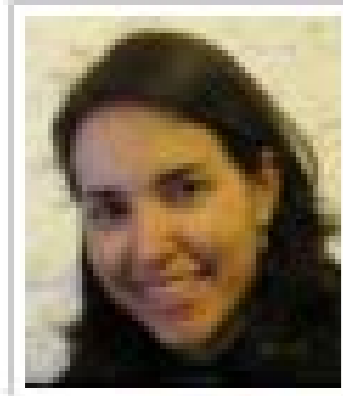
### MEMBERS



#### Francisco Javier Martínez de Pisón Ascacibar

Head of the EDMANS group. Associate professor at University of La Rioja. He is Industrial Engineer and PhD in Industrial Engineering by the University of La Rioja. His research activities focus on the use of soft computing, data mining and machine learning methods to solve real problems in several fields as industry, energy, agriculture and business.

email: [fmartin@unirioja.es](mailto:fmartin@unirioja.es)



#### Alpha Pernía Espinoza

Alpha is Associate Professor of Universidad de La Rioja (Spain) since 2012. She is Industrial Engineer and PhD in Industrial Engineer by the Universidad de La Rioja, and Electrical Engineer and Master in Industrial Automation and Instrumentation by the Universidad de Los Andes (Venezuela). She is officially member of the EDMANS group since 2006. Her research interests involve industrial processes modelling and optimization through Finite Element Analysis (FEM) and Computer Fluid Dynamics (CFD). In her free time she enjoys tennis and hiking.

email: [alpha.pernia@unirioja.es](mailto:alpha.pernia@unirioja.es)



#### Esteban Fraile García

Esteban is an Associate Professor at the Mechanical Engineering department of University of La Rioja. He belongs to the continuum mechanics area. He has a experience of 15 years in private companies in the civil engineering sector. His research is focused on sustainable buildings and on the development of new materials.

email: [esteban.fraile@unirioja.es](mailto:esteban.fraile@unirioja.es)



#### Eduardo Martínez de Pisón Ascacibar

Eduardo is an Associate Professor at the Mechanical Engineering department of University of La Rioja. He teaches in the continuum mechanics area. He researches on innovative materials with structural purposes.

email: [eduardo.mrtzdepison@unirioja.es](mailto:eduardo.mrtzdepison@unirioja.es)



#### Julio Fernández Cenicerós

Julio is Industrial Engineer and Master in Project Management by the University of La Rioja. In September 2009 he officially joined to the EDMANS group. In 2010, he was awarded with a FPI scholarship (2010-2014) by the University of La Rioja and under the supervision of Prof. Francisco Javier Martínez de Pisón Ascacibar. He is currently finishing his PhD in steel connections modeling by combining the finite element method (FEM) and predictive models based on soft-computing and data mining techniques. Outside research, he enjoys running.

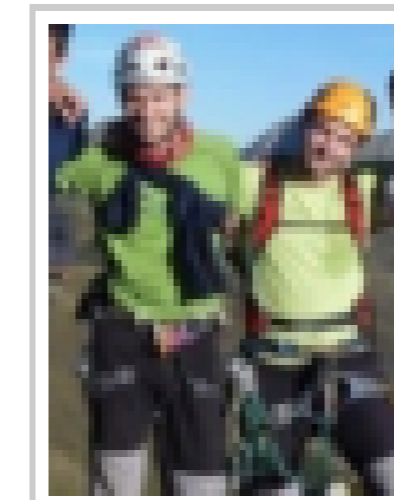
email: [julio.fernandezc@unirioja.es](mailto:julio.fernandezc@unirioja.es)



#### Fernando Antoñanzas Torres

Fernando is Industrial Engineering and Master in Renewable Energy and Energy Market by the Escuela de Organización Industrial (Madrid) and Product Engineering and Industrial Processes by the University of La Rioja. He has been officially in the EDMANS group since July 2012 with a FPI scholarship (2012-2016) by the University of La Rioja. He is doing his PhD in solar irradiation modeling under the supervision of Francisco Javier Martínez de Pisón Ascacibar (University of La Rioja) and Oscar Perpiñán Langueiro (Universidad Politécnica de Madrid). He is an enthusiast of rock climbing and traveling.

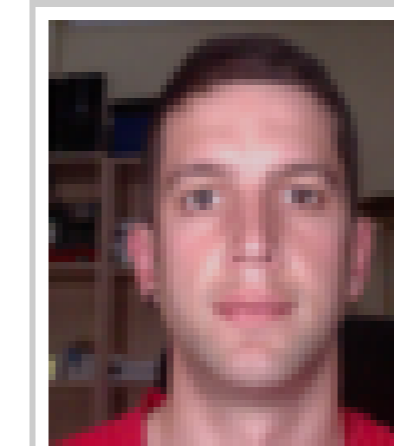
email: [antonanzas.fernando@gmail.com](mailto:antonanzas.fernando@gmail.com)



#### Andrés Sanz García

He is Industrial Engineer and PhD in Industrial Engineering by the University of La Rioja. He has performed as project manager in different engineering and consulting companies for ten years. In September 2010, he joined EDMANS working towards the application of Data Mining. He is actually in a fellowship in Japan while doing his second PostDoc in bioprinting and NGS techniques at Finland. He loves practising any kind of sport, climbing and running in particular.

email: [andres.sanz@unirioja.es](mailto:andres.sanz@unirioja.es)



#### Rubén Urraca Valle

Rubén is Industrial Engineer and Master in Project Management by the University of La Rioja. He has been officially in the EDMANS group since September 2014 with a FPU scholarship (2014-2016) by the Ministry of Education, Culture and Sports of Spain. In his spare time he enjoys cycling.

email: [ruben.urraca@unirioja.es](mailto:ruben.urraca@unirioja.es)



#### Javier Antoñanzas Torres

Javier is an Industrial Engineer by the University of La Rioja and Master in Renewable Energy and Energetic Efficiency by the University of Zaragoza. He joined the EDMANS group in September 2014 with a FPI scholarship from the University of La Rioja (2014-2016). His passion is climbing in the mountains.

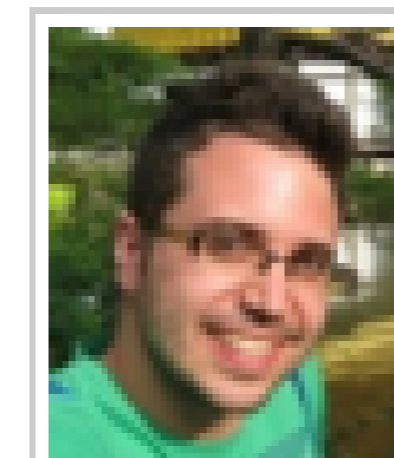
email: [antonanzas.javier@gmail.com](mailto:antonanzas.javier@gmail.com)



#### Javier Ferreiro

Javier is an Associate Professor at the Mechanical Engineering department of University of La Rioja. He has an experience of 13 years in private engineering companies in the construction sector. His research focuses on sustainable buildings analyzing costs, environmental and social impacts.

email: [javier.ferreiro@unirioja.es](mailto:javier.ferreiro@unirioja.es)



#### Enrique Sodupe Ortega

PhD student. Enrique is an Industrial Engineer by University of La Rioja. He has been working in EDMANS group with two different scholarships: "Beca de colaboración de la Universidad de La Rioja (2011-2012)" and "Beca de colaboración del MEC (2012-2013)". He is granted by a FPI-CAR fellowship (2015-2016) by the Government of La Rioja. He is mainly interested in CFD simulation and 3D printing. In his spare time he enjoys running and traveling.

email: [esodupe@gmail.com](mailto:esodupe@gmail.com)



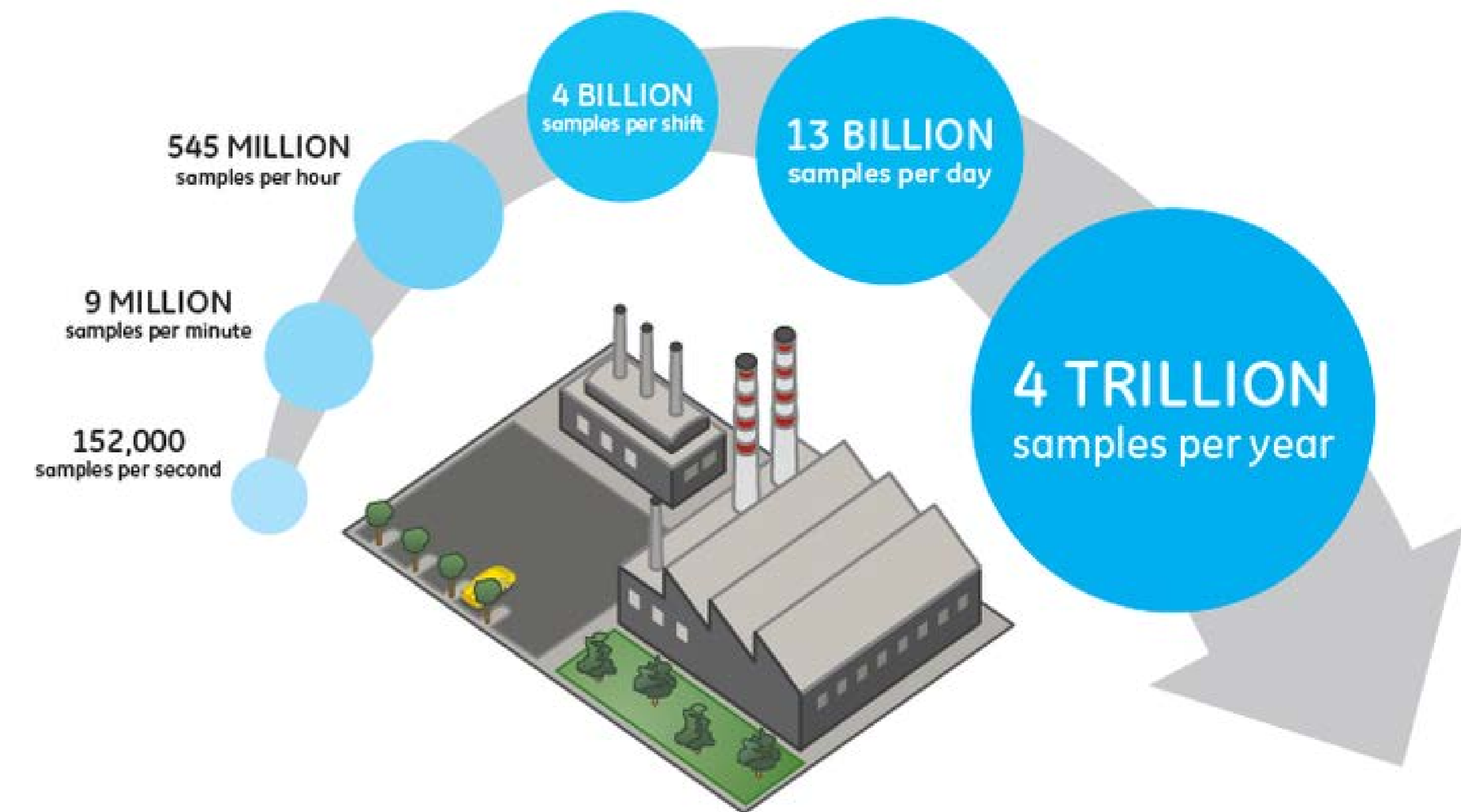
# Big Data (BD)

La Información externa e interna pueden Generar Valor a la Organización....



# Industrial Big Data (IBD) vs BD

- Gran parte de los datos provienen de **equipos y procesos automatizados**.
- En IBD, los datos son más estructurados, correlacionados y organizados en el tiempo.
- Se trabaja habitualmente con **Históricos** del proceso industrial (pasado).
- **Objetivo: extraer conocimiento oculto y relevante del proceso industrial que permita entenderlo mejor y utilizar el mismo para la mejora en la toma de decisiones.**
- **Resultados:** mejora de la calidad, reducción de costes, reducción de tiempos, reducción de tasa de fallos, más y mejor conocimiento del proceso, más flexibilidad, mejor control del proceso, más seguridad, etc.



“The Rise of Industrial Big Data”. GE Intelligent Platforms, 2011.  
[http://leadwise.mediadroit.com/files/19174\\_the\\_rise\\_of\\_industrial\\_big\\_data\\_wp\\_gft834.pdf](http://leadwise.mediadroit.com/files/19174_the_rise_of_industrial_big_data_wp_gft834.pdf)

# IBD en el 2020

Investments  
40 billion Euros per year  
2020

Source: Germany  
Trade & Invest  
(GTAI)

more than  
**80%**  
Digital Economy  
by 2021



**13%**  
cost reduction

**18%**  
efficiency increase

# Técnicas y Métodos en IBD

- **Big Data Architectures:** MapReduce, Hadoop, Spark, ...
- DataWarehouse methods
- Visual Data Mining
- Deep Learning and Neural Networks (DNN, Convolutional NN, ...)
- Bayesian Methods
- Support Vector Machines
- Tree-based Methods
- Clustering
- Ensemble Methods: Bagging, Adaboost, Gradient Boosting Machines,...
- Instance Based Learning
- Linear and Non-Linear Mapping
- Advanced Multivariate Statistics
- Advanced
- Bio-inspired computing and Metaheuristic optimization algorithms: Genetic Algorithms, Genetic programming, Particle Swarm Optimization, Ant Colony, Bees Alg., Swarm Int., ...
- Fuzzy Methods
- Hybrid Artificial Intelligent Systems
- Advanced Database methods
- ...



# Caso Práctico I: Sensor Software

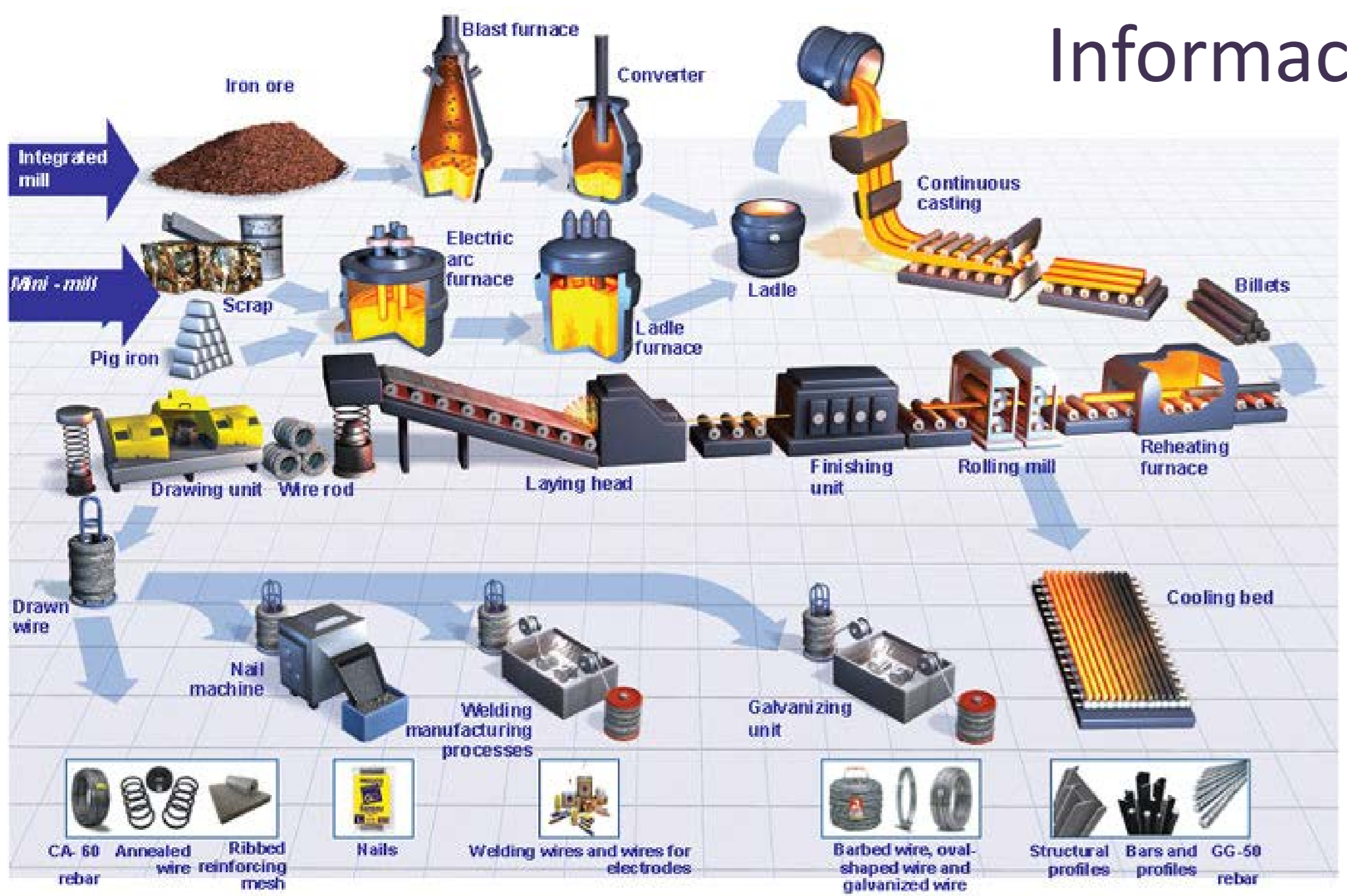


ArcelorMittal

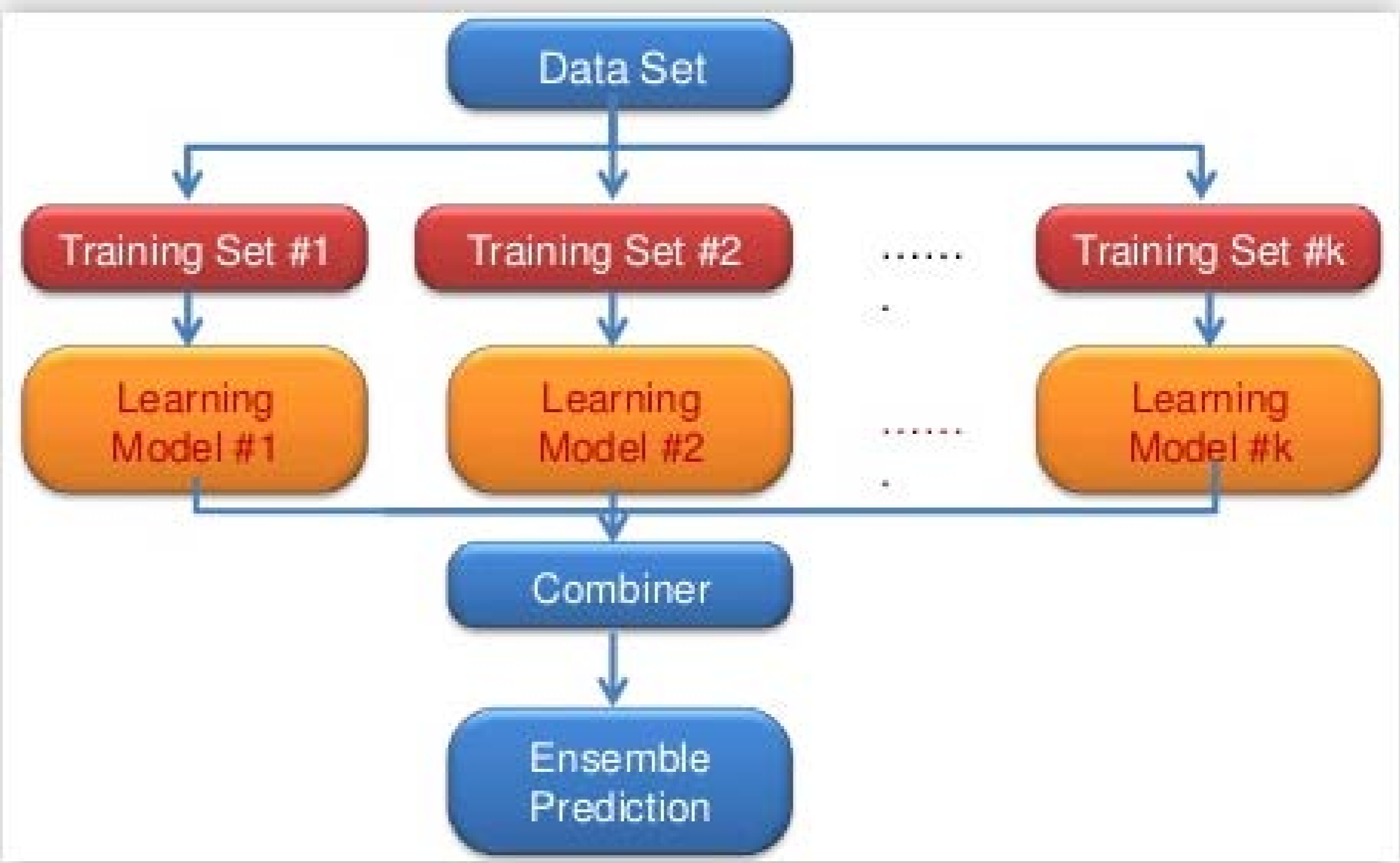
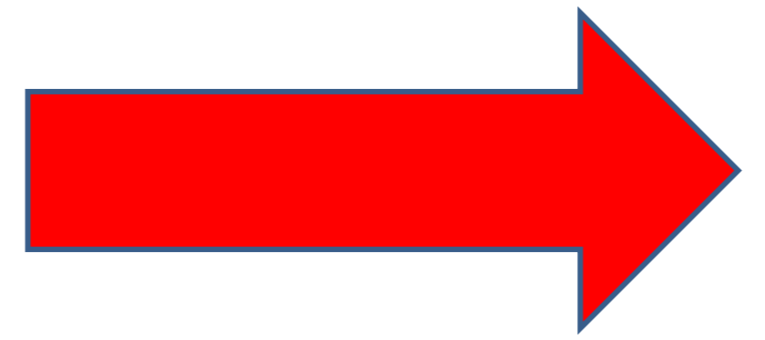
- Problemática: una compañía fabricante de carrocerías te devuelve el barco si **3 bobinas no cumplen** las características mecánicas del acero pedidas (límite de elasticidad, resistencia a la tracción, diagrama tensión-deformación, etc.) debido a que la producción se para si se atasca o rompe la prensa...
- No existen sensores que midan esos parámetros en cada metro de bobina de acero.
- Disponemos de miles de datos del proceso y de cada bobina se extrae una probeta y se realiza un ensayo de tracción para determinar sus características mecánicas....



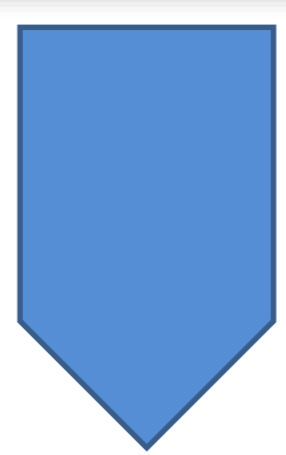
# Solución: Sensor Software con "Ensemble Models"



Información del Proceso Industrial



Caracterización Producto



Ejemplo de Predicción del Límite Elástico cada 100m

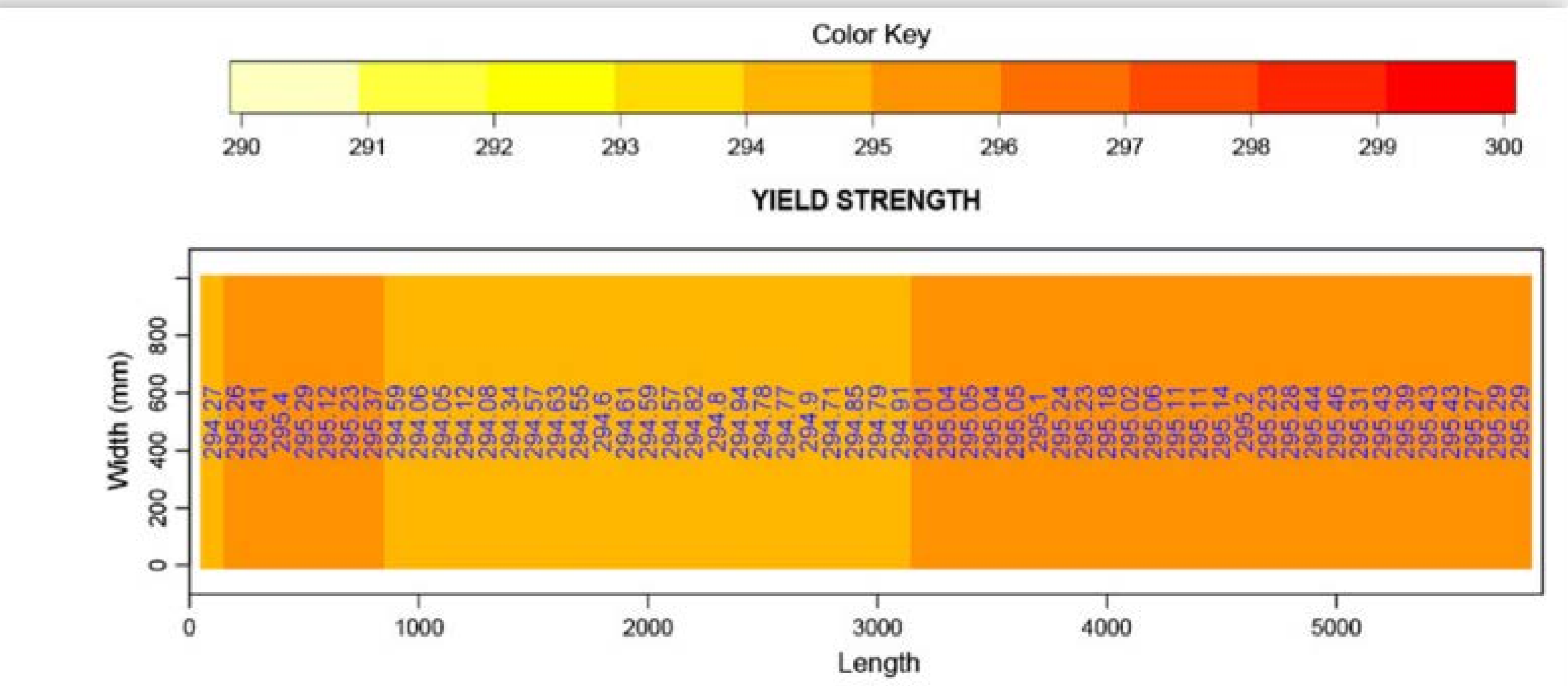
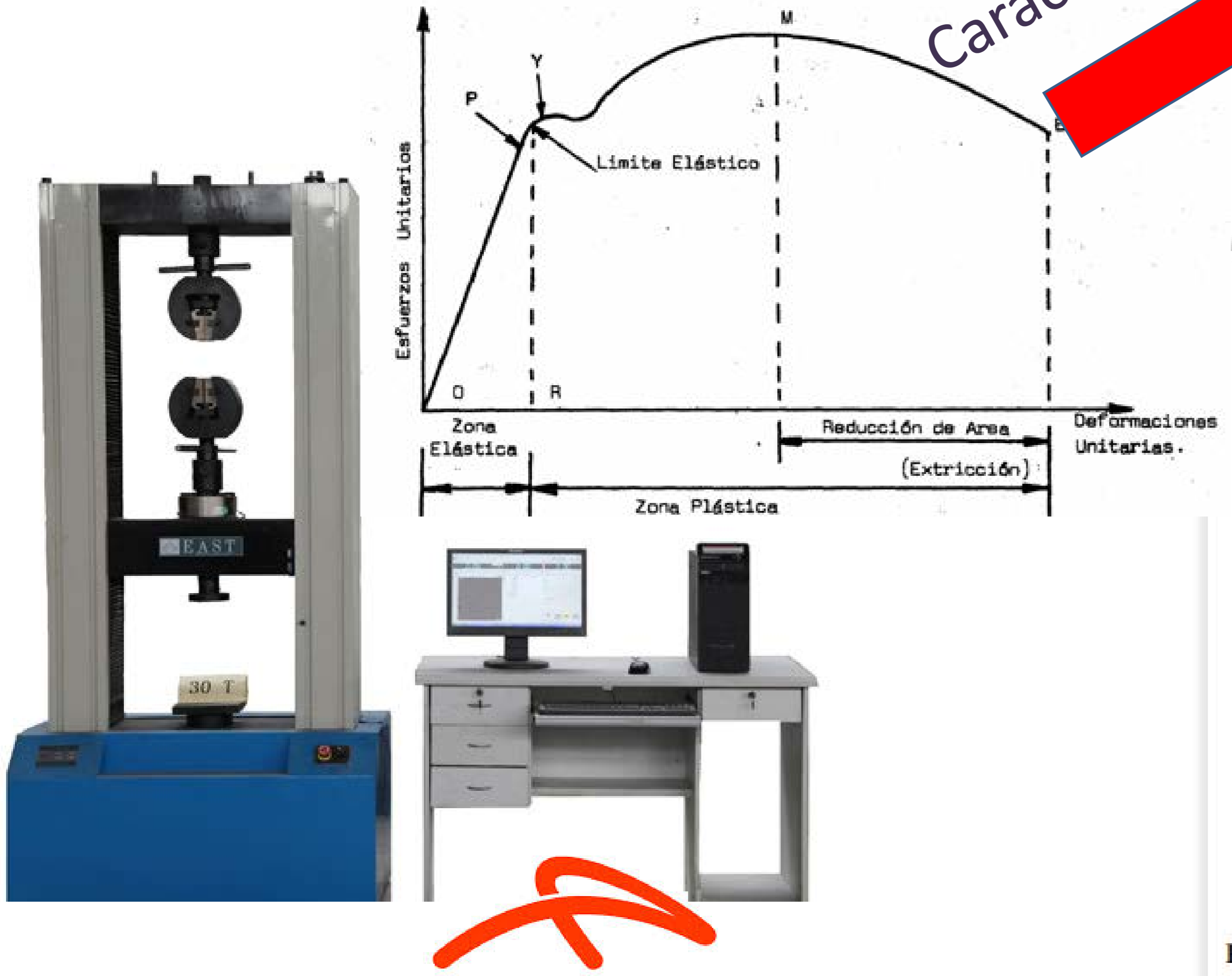


Fig. 8 Estimation of mechanical properties throughout the coil as proposed by the model

# Caso Práctico II: Reducción de Defectos



ArcelorMittal

- Problemática: se encuentran defectos superficiales en el proceso de galvanizado de las bobinas de acero.
- Se disponen de miles de variables del proceso y de imágenes obtenidas con cámaras lineales de los defectos.



Defect density 3D

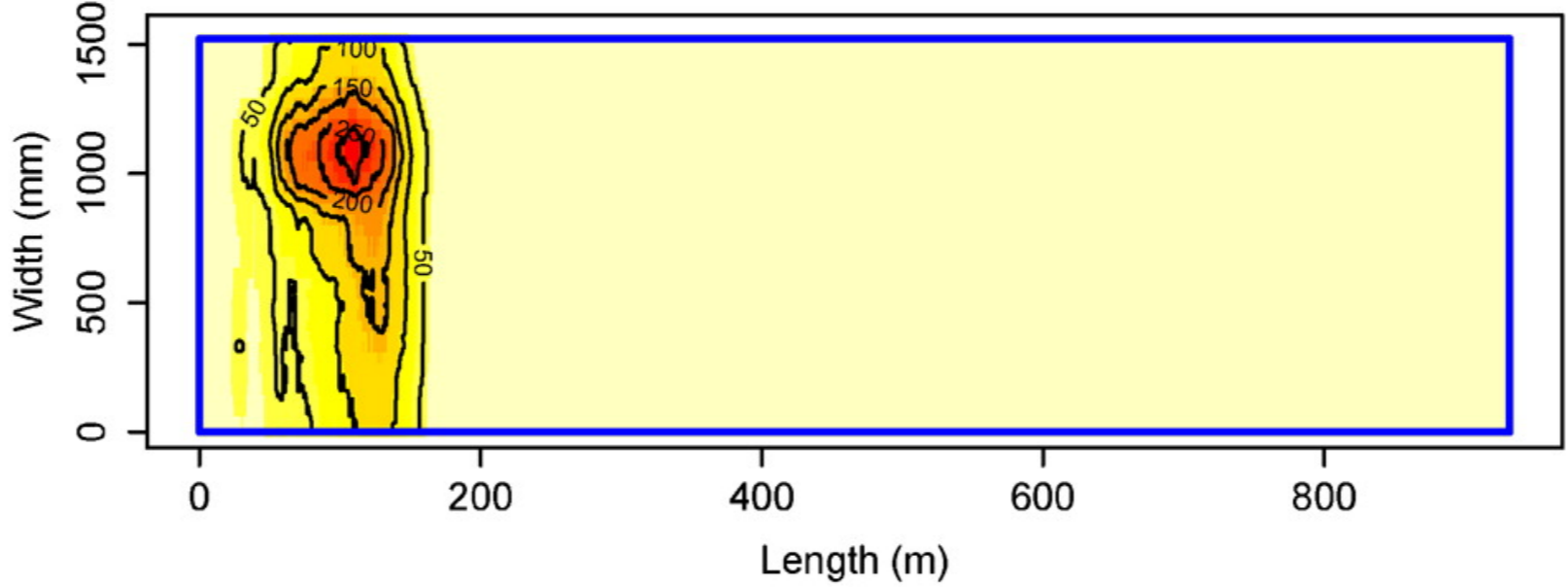
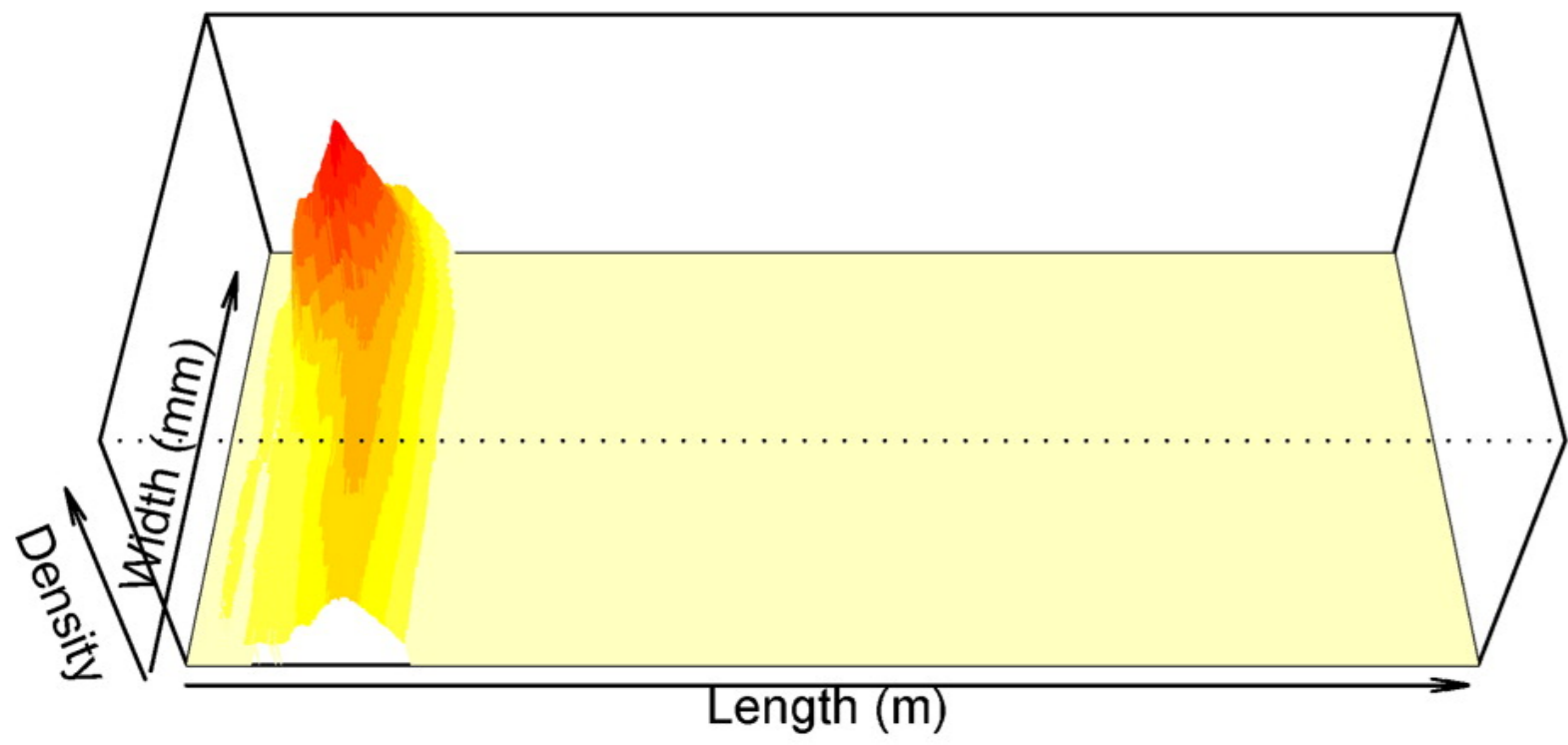
Defect density 2D

Color Key



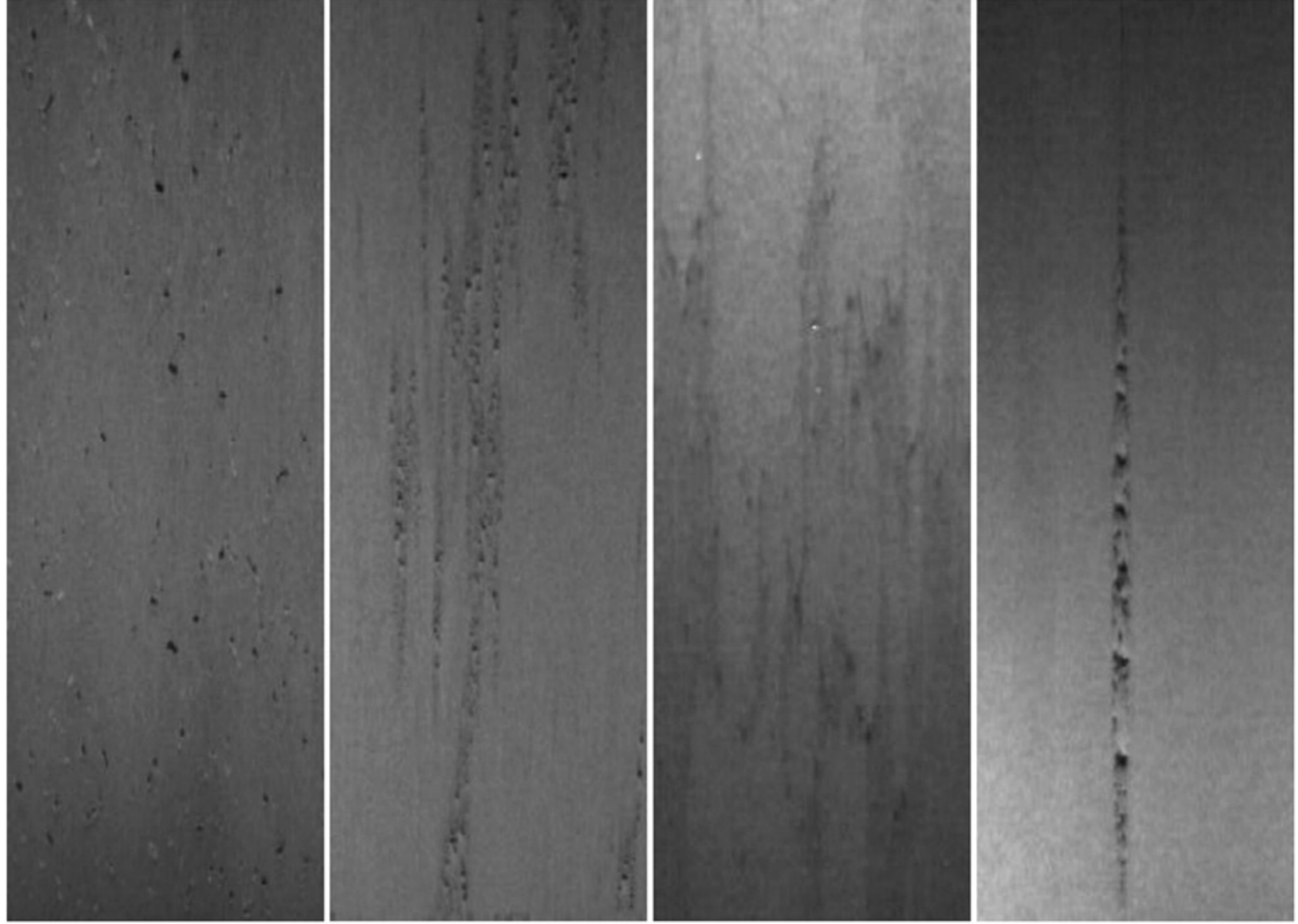
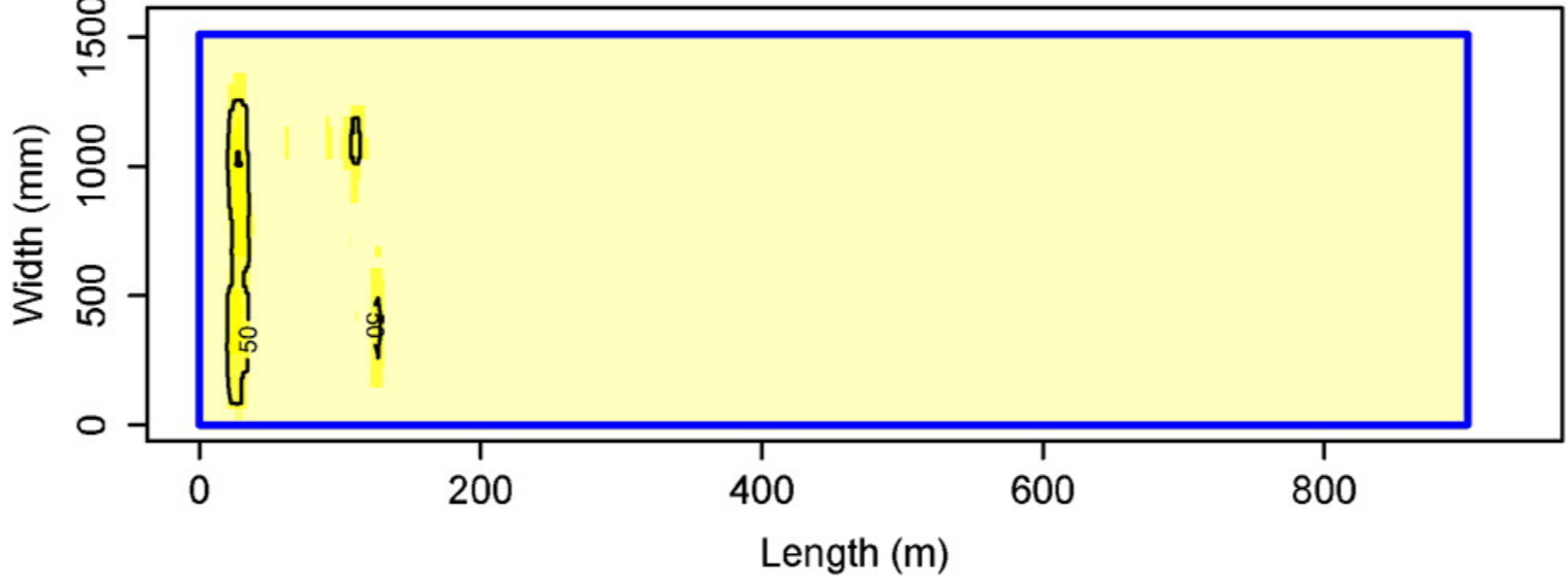
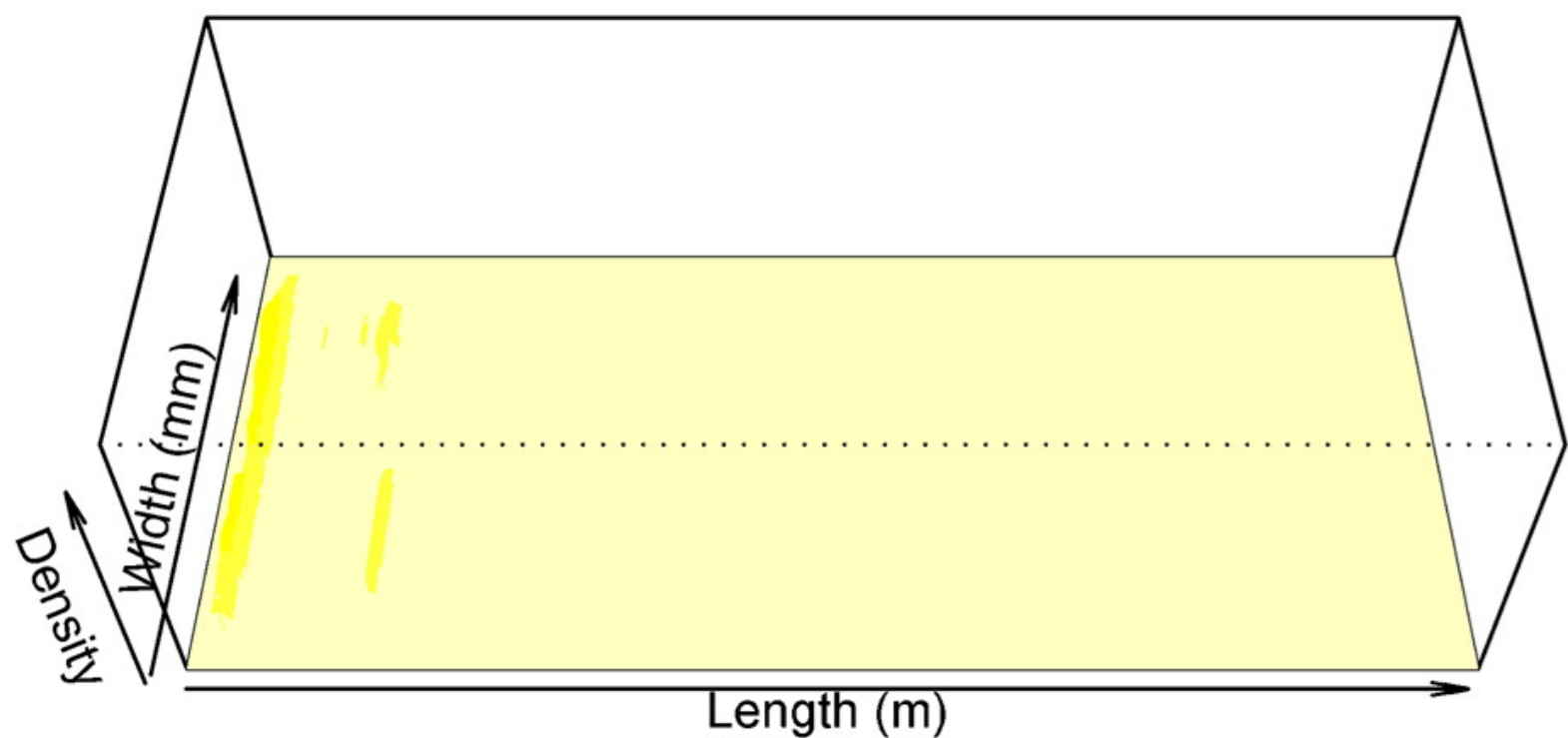
HSM (Side 0)

HSM (Side 0)



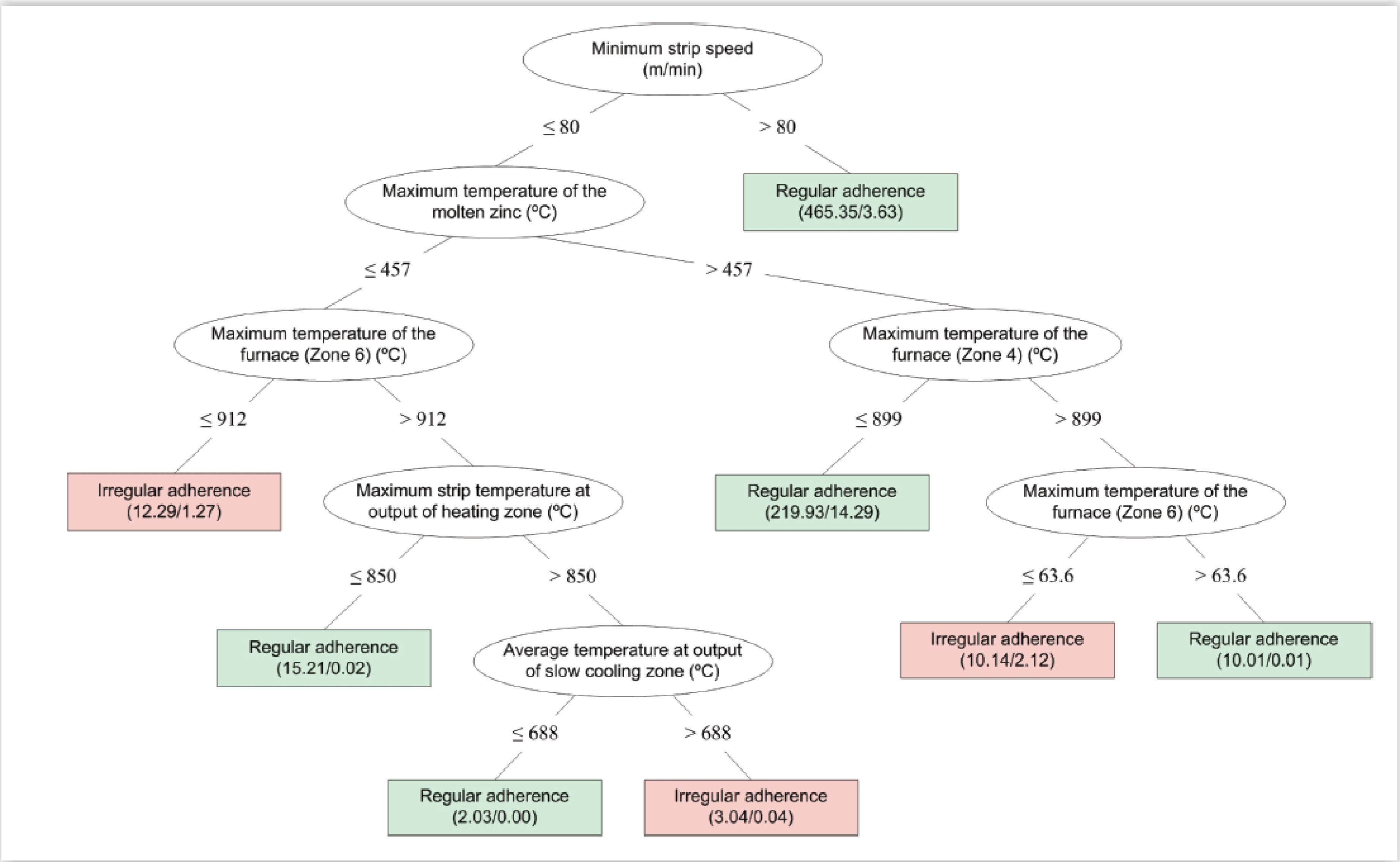
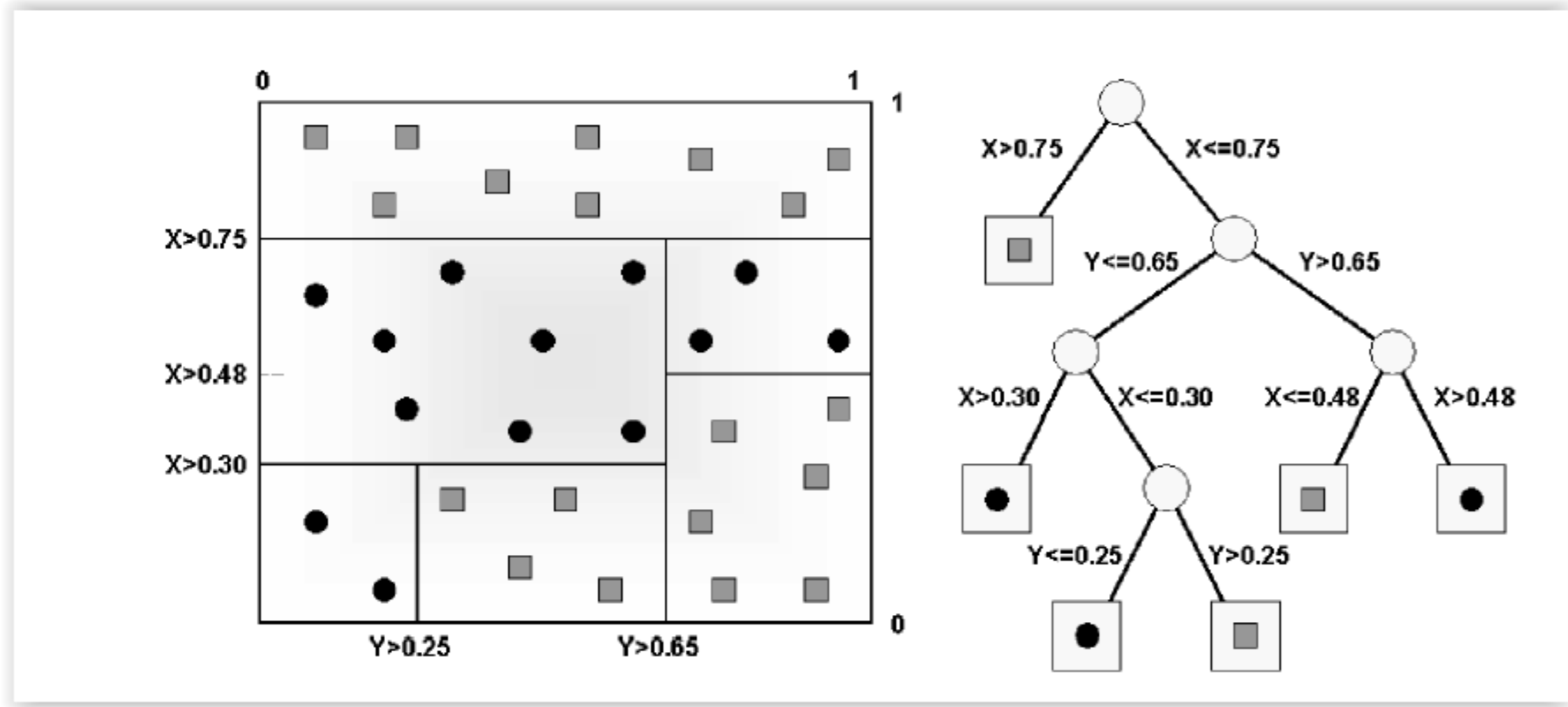
PICKLING LINE (Side 0)

PICKLING LINE (Side 0)



a) C7 scale b) C3 scale c) Secondary scale d) C5 scale

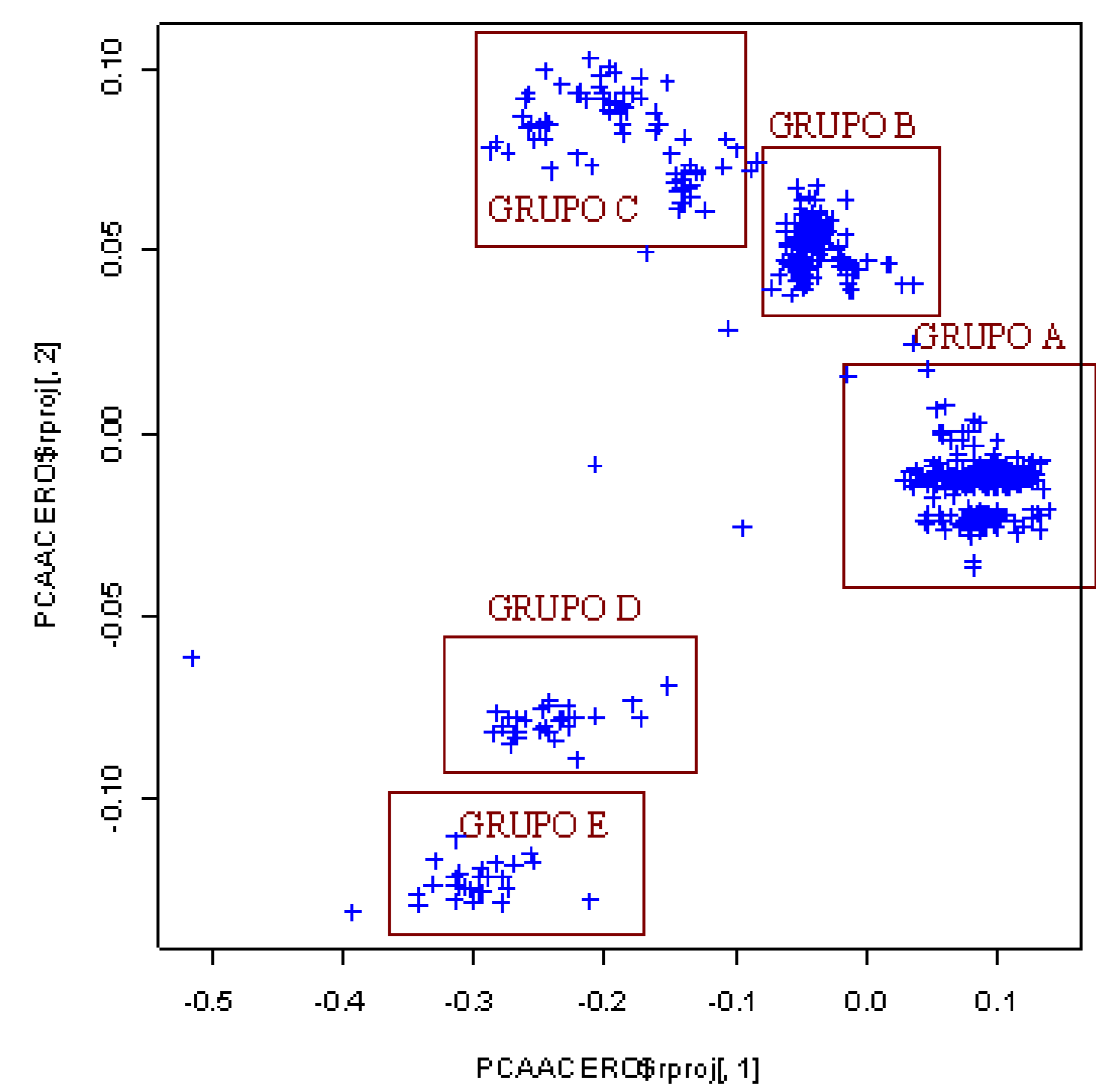
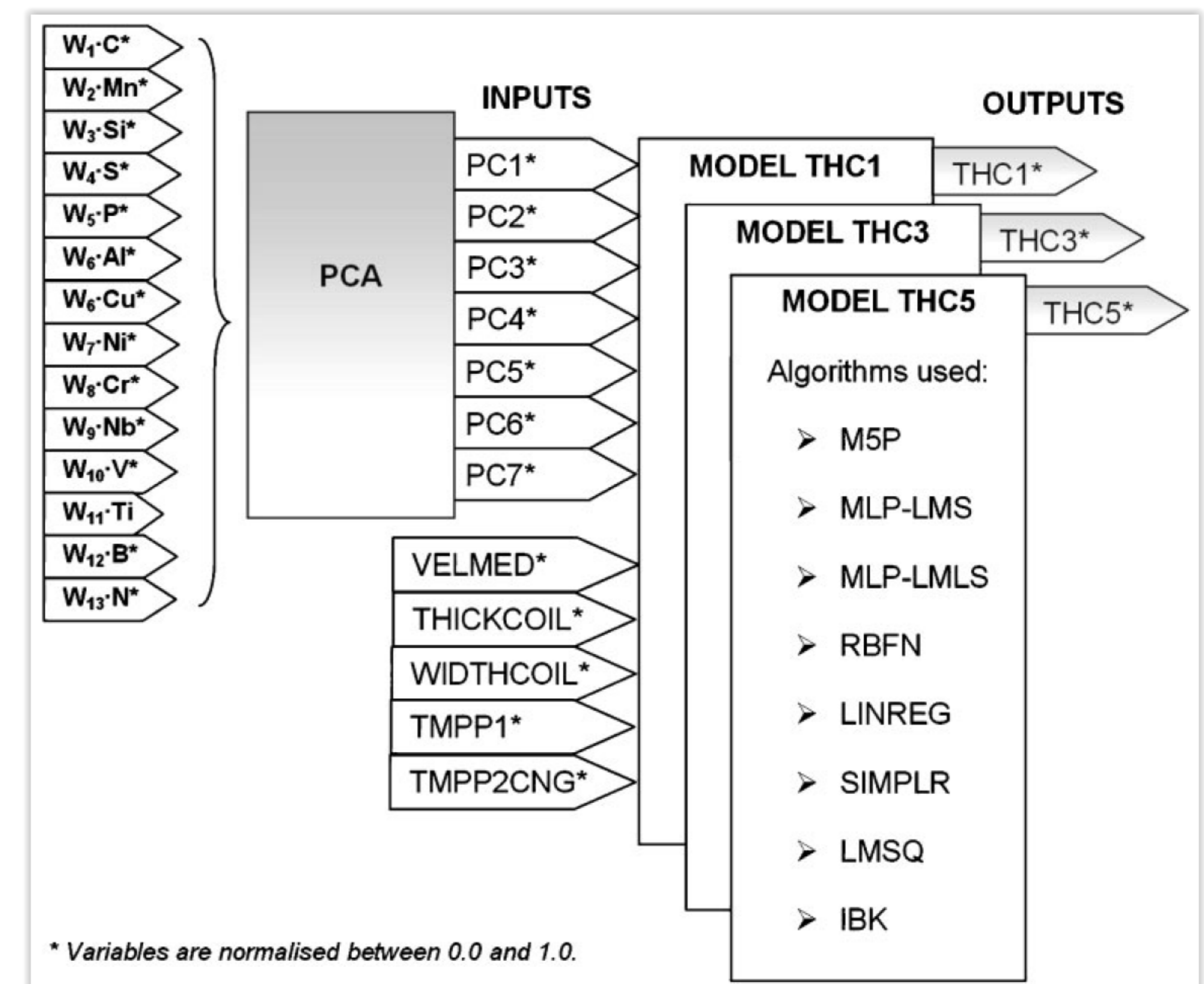
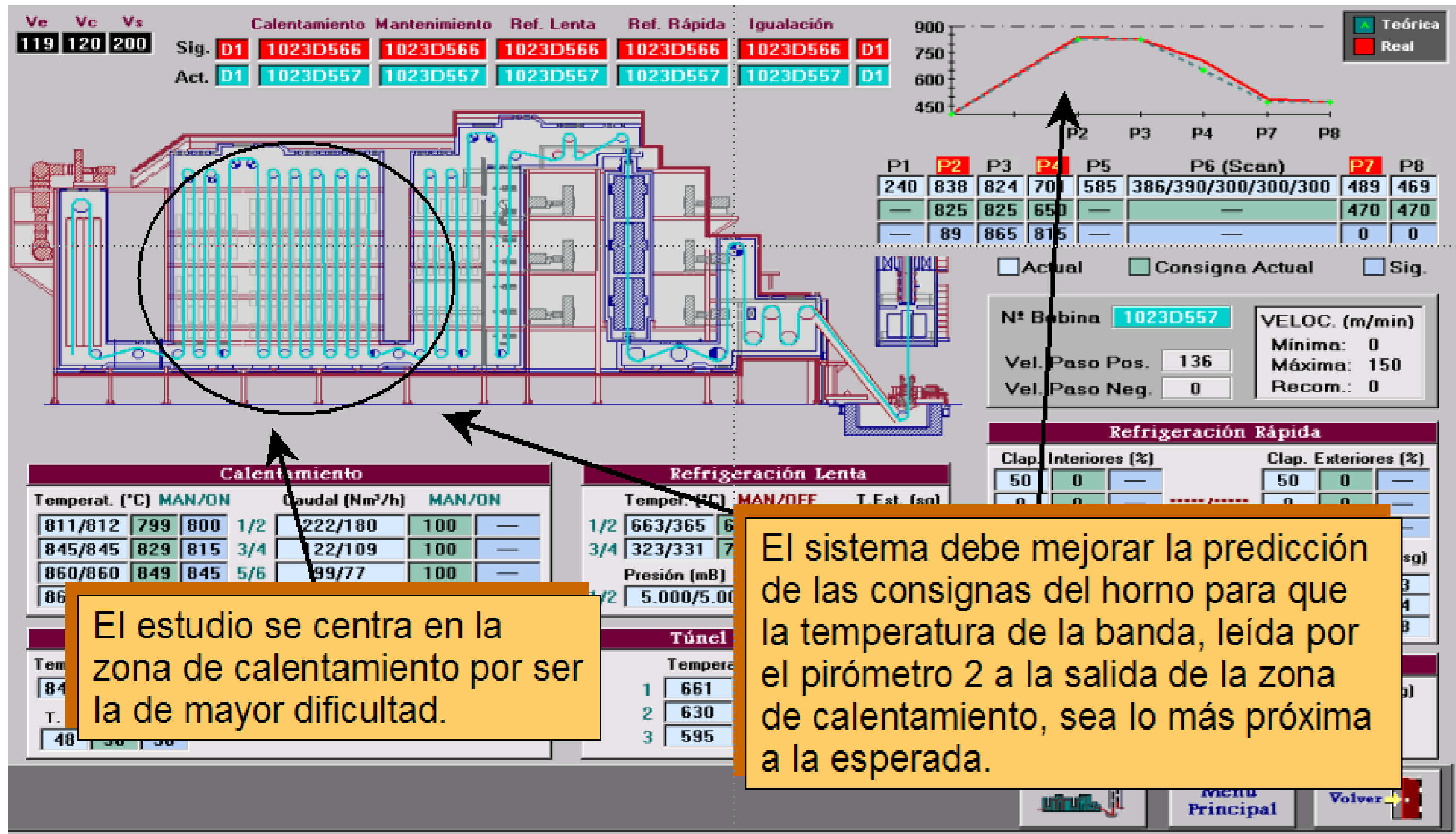
# Primer Paso: Análisis del Proceso y Extracción de Conocimiento Oculto con Árboles de Decisión



GRUPOS DE VARIABLES	
Nombre	Descripción
CODE	Código de la bobina.
CON_H2	Concentración de H <sub>2</sub> en el aire presente en esa zona.
CON_O2	Concentración de O <sub>2</sub> en el aire presente en esa zona.
TMP_PR	Temperatura del punto de rocío en esa zona.
Zxx_TMP	Temperatura de la zona xx.
TMP_Pxx	Temperatura de la banda, medida con el pirómetro xx, en cada zona del proceso.
ESP, ANCH	Espesor y anchura de la banda de acero.
SPD	Velocidad de la banda de acero.
POT_TMP	Temperatura del baño de zinc.
CMP_BAÑO	Composición química del baño de zinc. Porcentaje de los elementos químicos más relevantes del baño de zinc.
CMP_ACERO	Composición química del Acero. Porcentaje de los elementos químicos más relevantes: Fe, Mn, Al, Ni, etc.
ADHERENCIA	Si la adherencia del recubrimiento de zinc fue correcta o no (Valores: TIPO-A o TIPO-B)

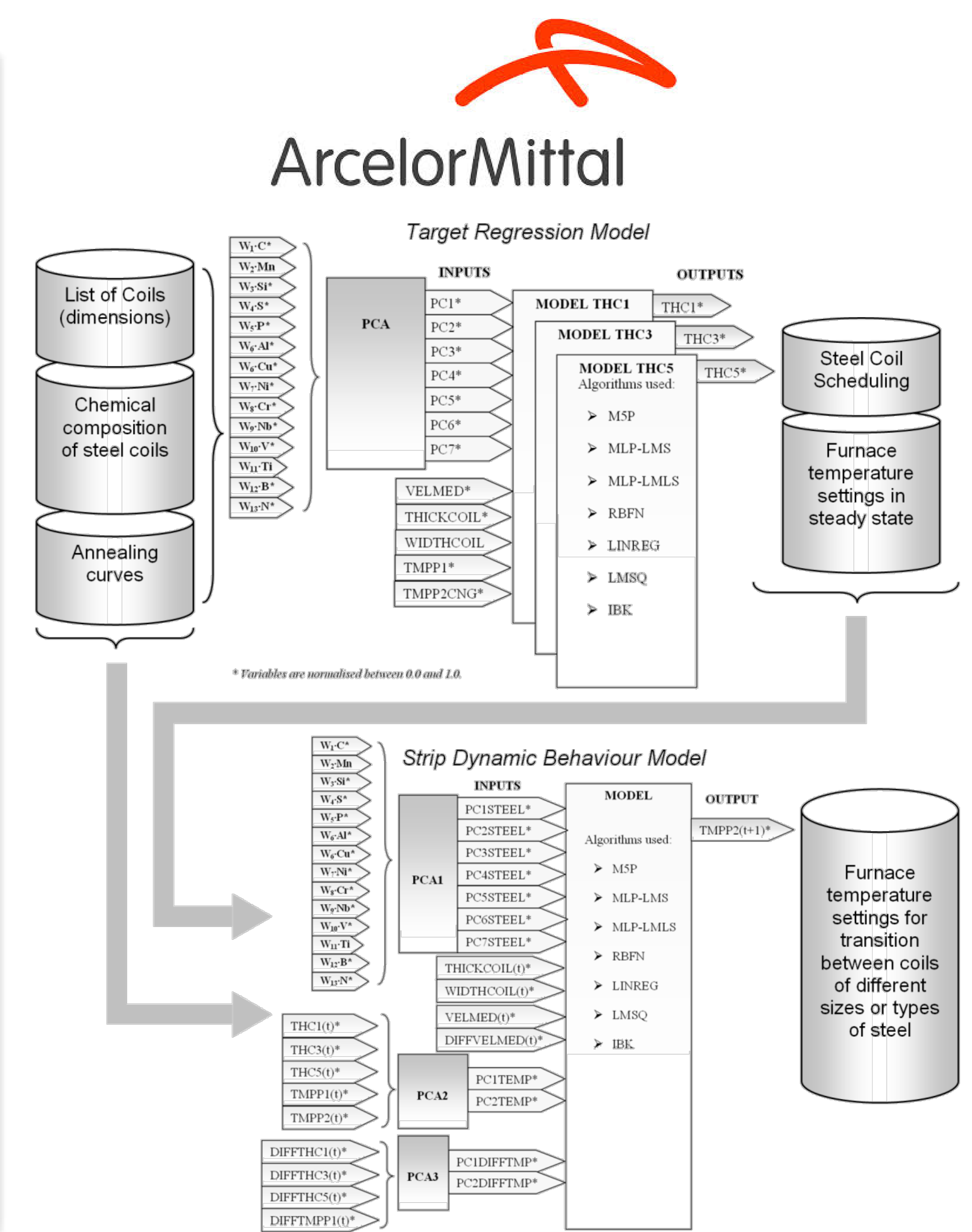
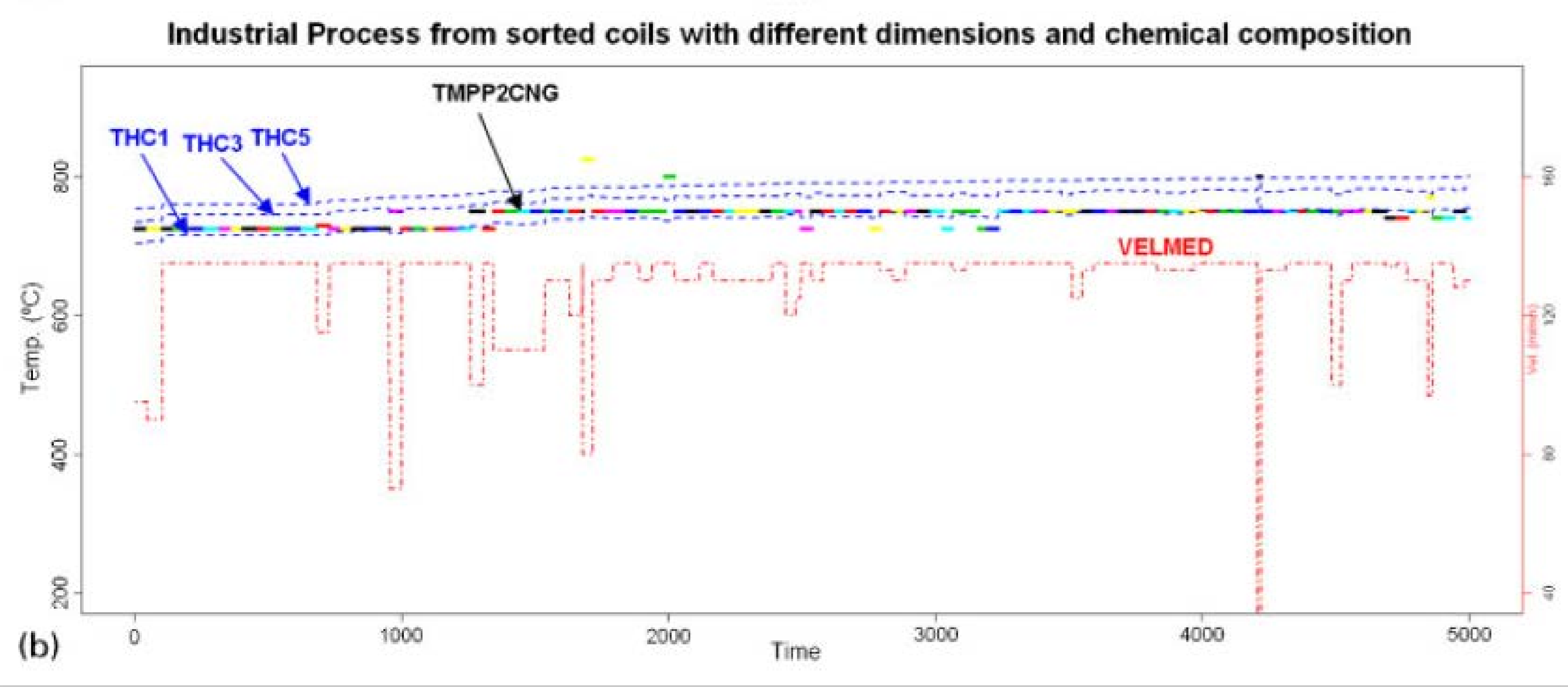
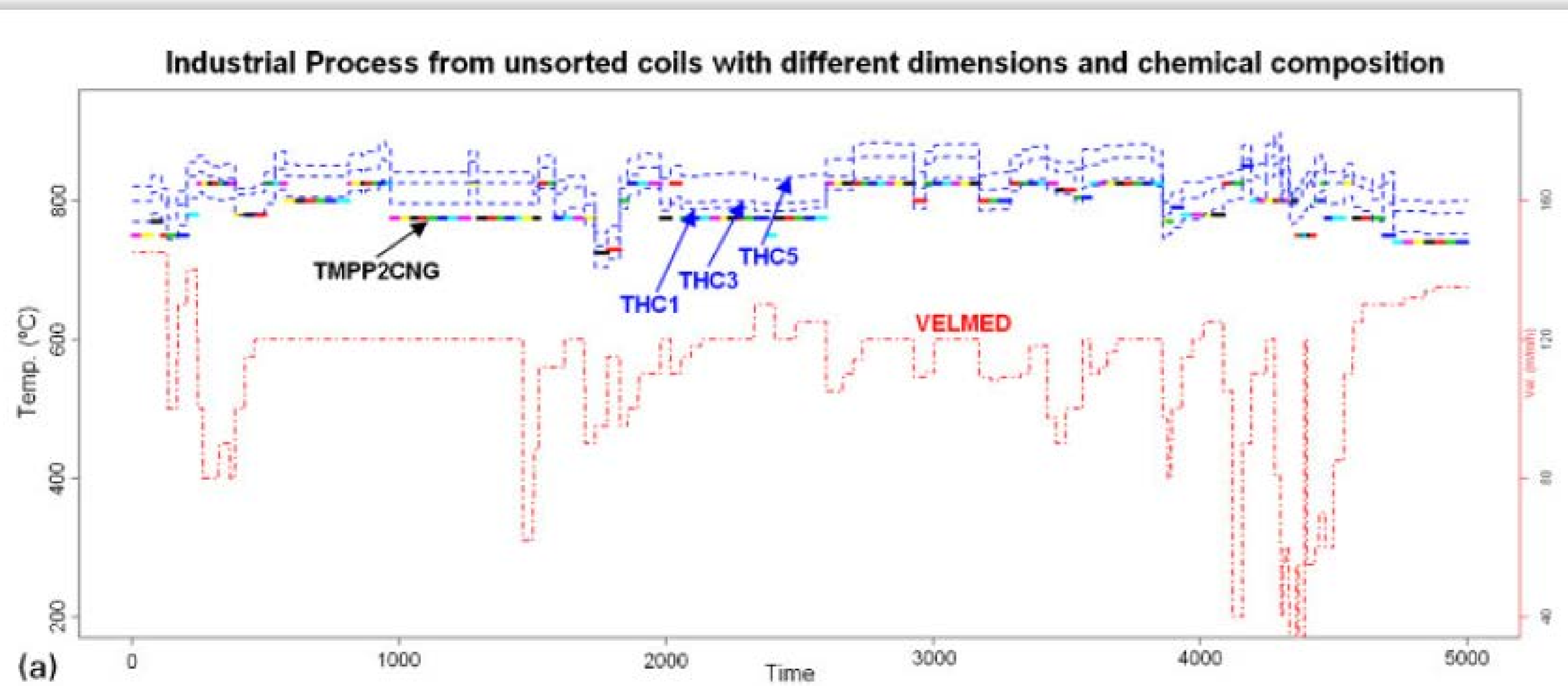
*EU Projects: RFS-CR-04023 & RFS-CR-04043*

# Segundo Paso: Desarrollo de modelos de control del horno que aprenden cuando se ha procesado correctamente las bobinas en modo “automático” o “manual”.

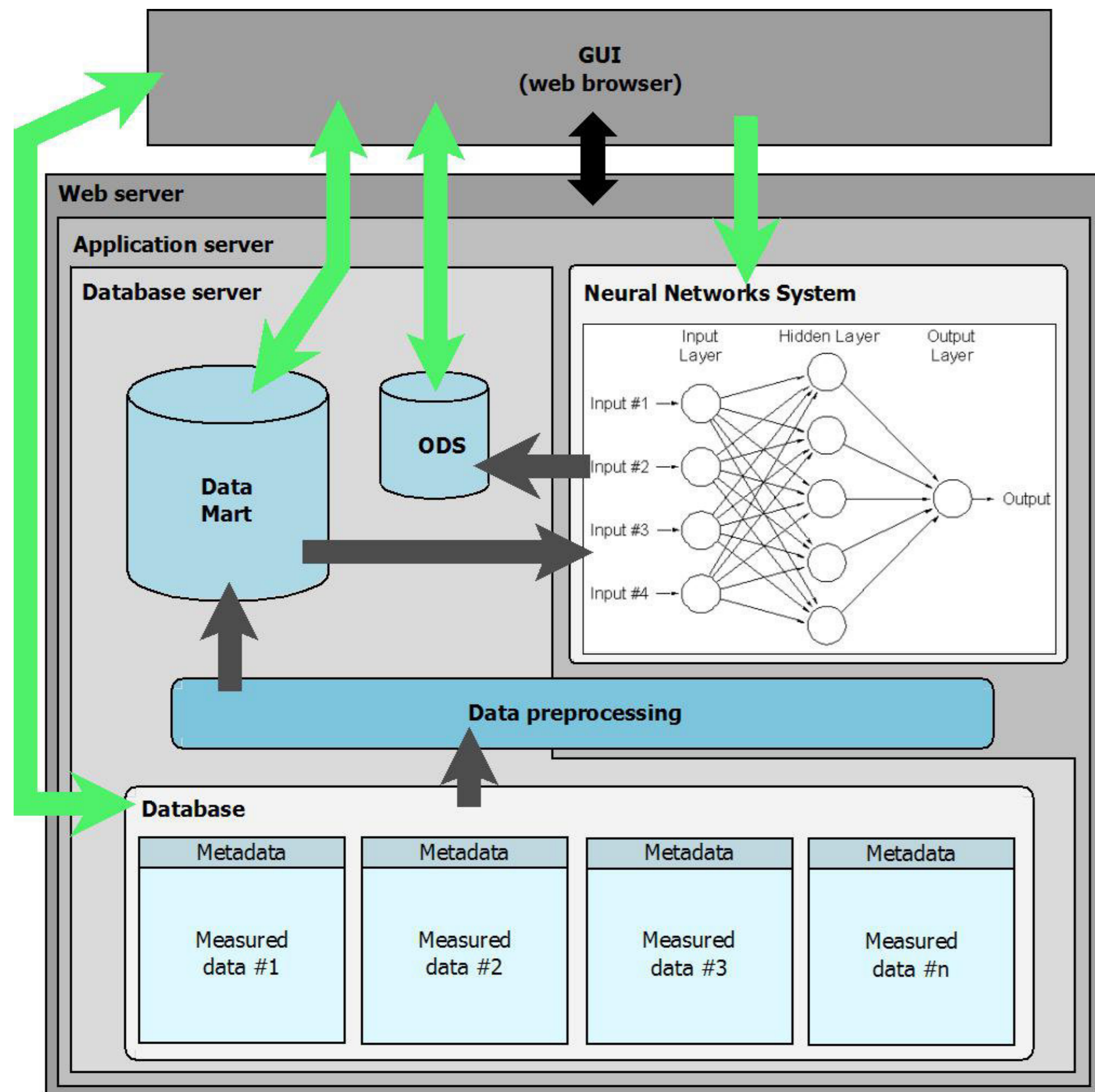
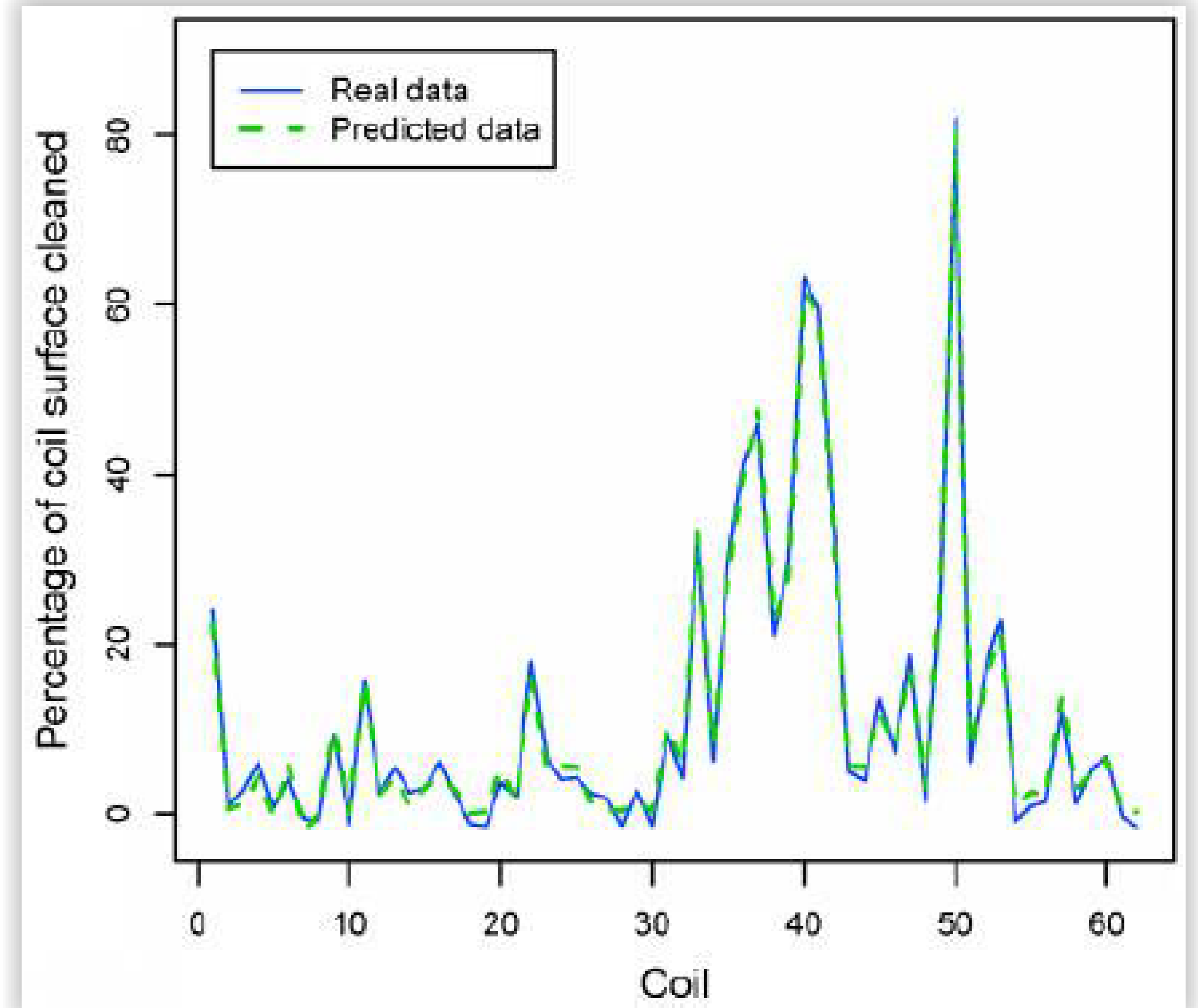
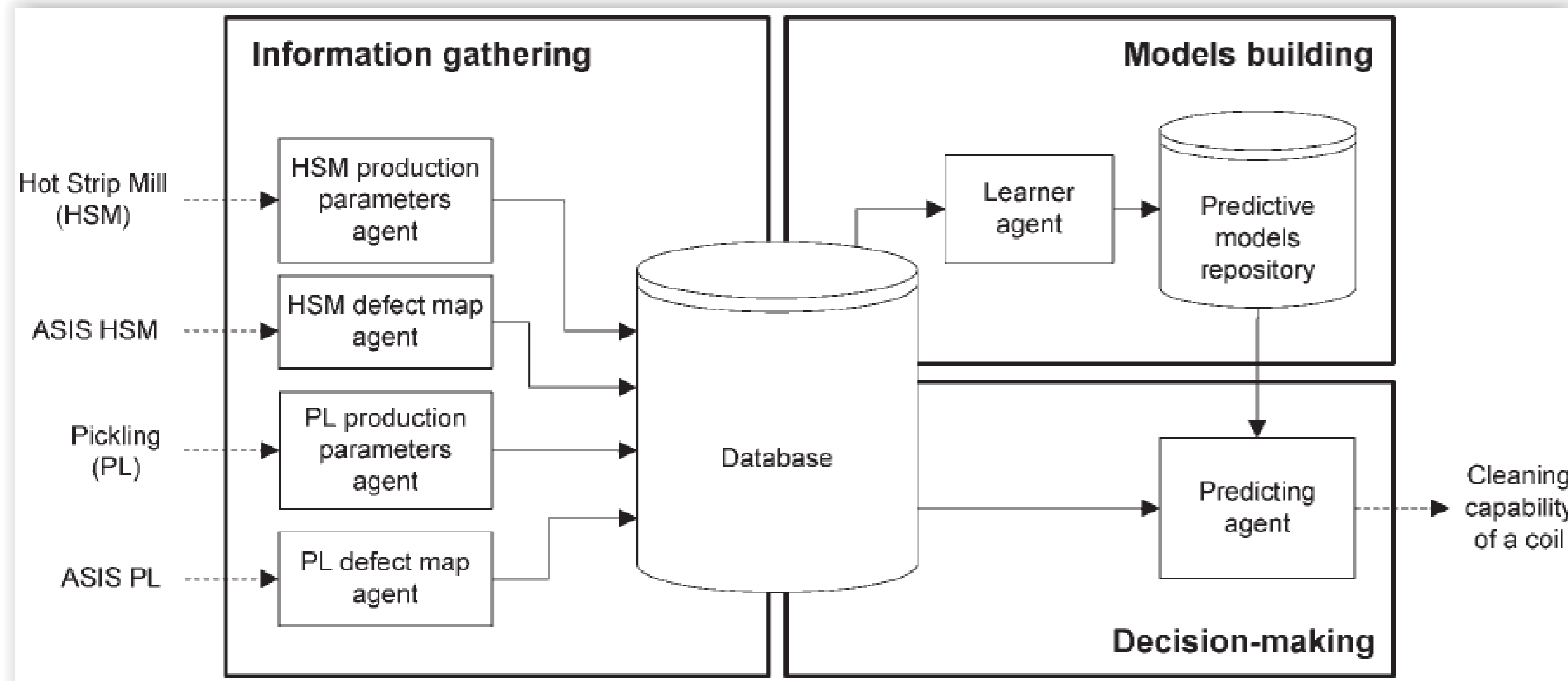


EU Projects: RFS-CR-04023 & RFS-CR-04043 & RFS-CR-03012

# Tercer Paso: Mediante los modelos desarrollados y la información proveniente de cada bobina optimizar con Algoritmos Genéticos el “scheduling” de las bobinas para poder reducir los “saltos”, reducir defectos.



# Cuarto Paso: desarrollo de un sistema inteligente multiagente de apoyo a la toma de decisiones capaz de detectar las bobinas que puedan tener problemas en el galvanizado a partir de cómo se ha comportado en procesos anteriores



EU Project:  
RFSR-PR-2006-00036

# Caso Práctico III: Mejora del Control de Calidad

- Objetivo: metodología para crear un modelo predictivo que permita determinar la probabilidad de que una pieza sea defectuosa a partir de una muestra de piezas que han pasado por el control de calidad.
- **Estimar probabilidad de ser defectuosas** en 1.183.748 piezas.
- A partir de otras 1.183.747 piezas de las cuales 6.879 (0,581%) defectuosas.
- **4.268 variables** en la BD con valores numéricos (dimensiones, espesores, etc.), tiempos de entrada y permanencia en cada estación, alarmas, estados, etc.
- **52 estaciones** de fabricación y ensamblado.



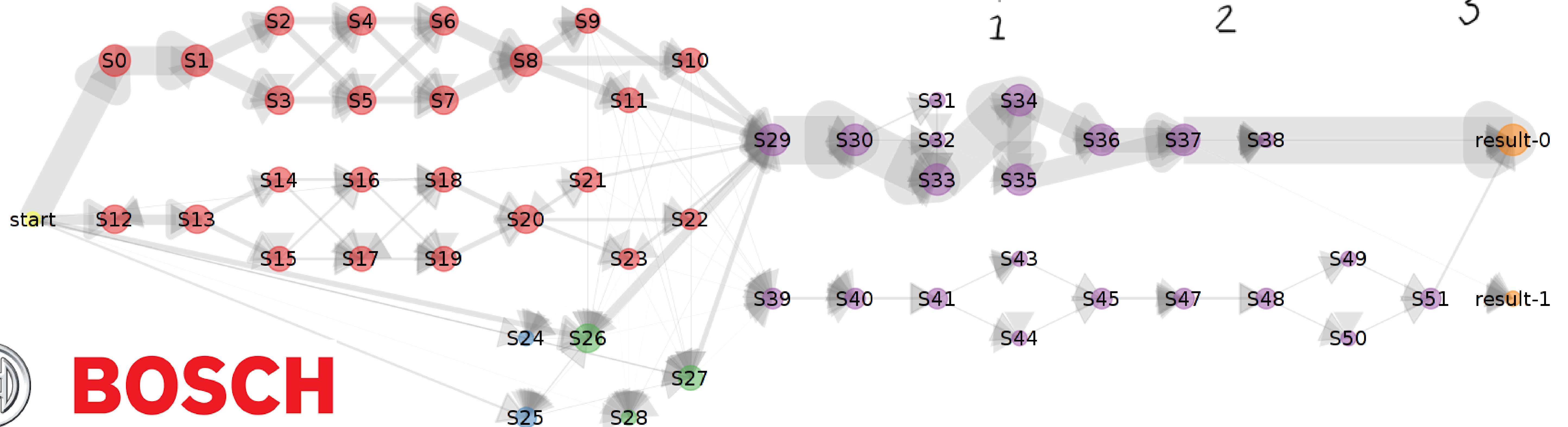
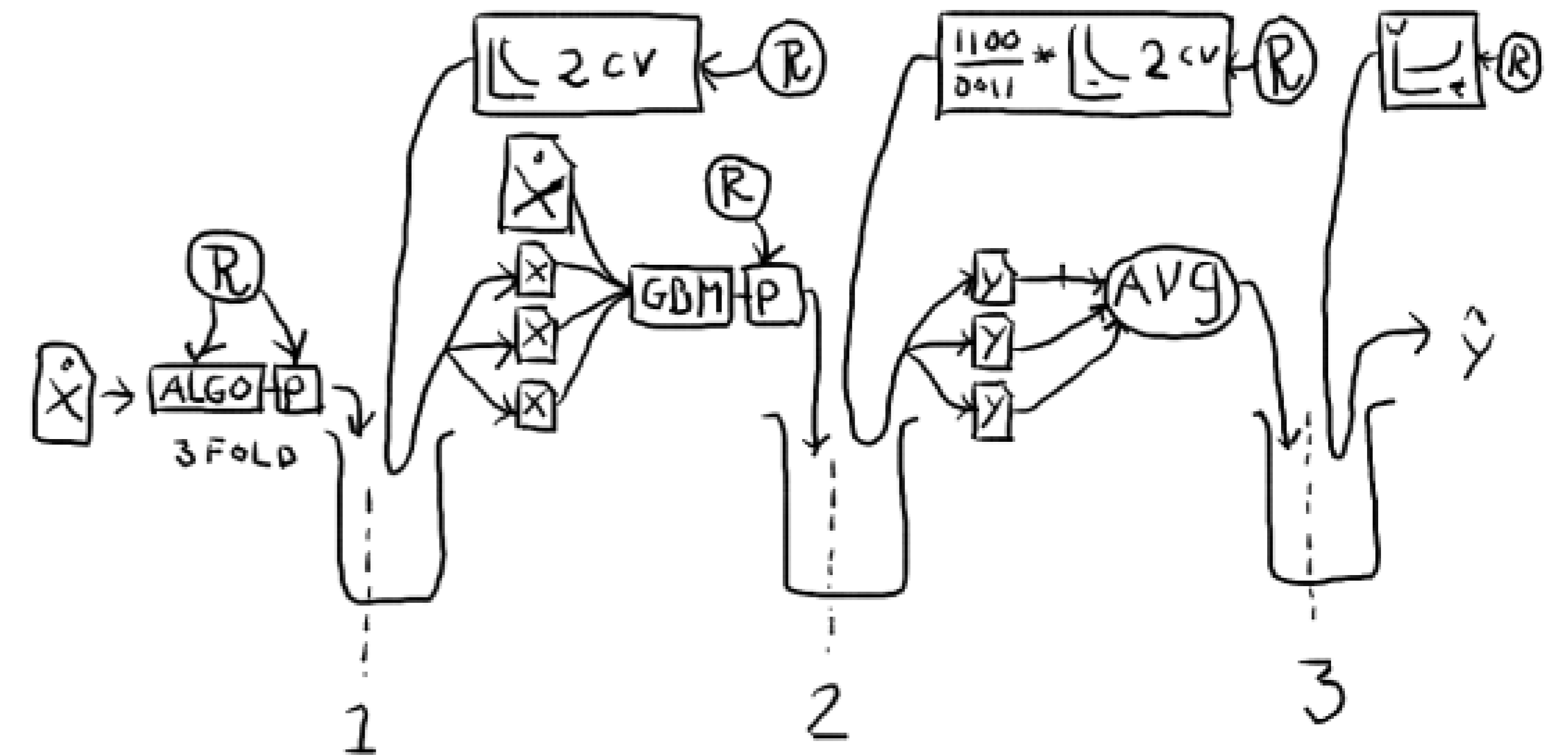
**BOSCH**



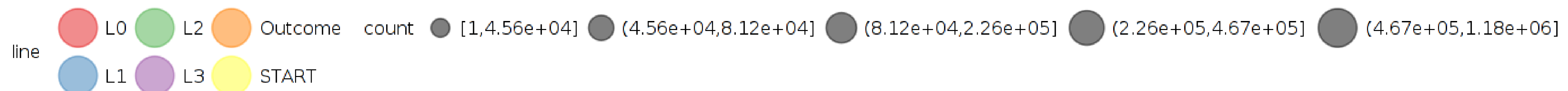


- Algoritmos capaces de dibujar solos el diagrama de flujo de las piezas fabricadas.
- Feature Selection and Feature Engineering:** Creación de nuevas variables que muestran diferencias de tiempo entre estaciones clave, orden en el tiempo en estaciones determinadas, media de fallos por unidad de tiempo, etc.
- Desarrollo de **modelos “ensemble” de tres niveles** basados fundamentalmente en Gradient Boosting Machines capaces de estimar la probabilidad con un Coeficiente de Correlación de Matthews de 0.5240.

Model	Submission Date	Public MCC	Private MCC	Rank
First XGB to check row order & time features	29 Sep 2016	0.4485	0.4606	95
XGB after paropt and more features	08 Oct 2016	0.4951	0.4968	24
L2 ensemble	18 Oct 2016	0.5071	0.5092	7
XGB with more time based features	26 Oct 2016	0.5010	0.5061	12
<b>L3 ensemble</b>	<b>30 Oct 2016</b>	<b>0.5165</b>	<b>0.5182</b>	<b>2</b>
Best XGB	11 Nov 2016	0.5119	0.5132	4
<b>Best selected L3 ensemble</b>	<b>11 Nov 2016</b>	<b>0.5243</b>	<b>0.5240</b>	<b>1</b>

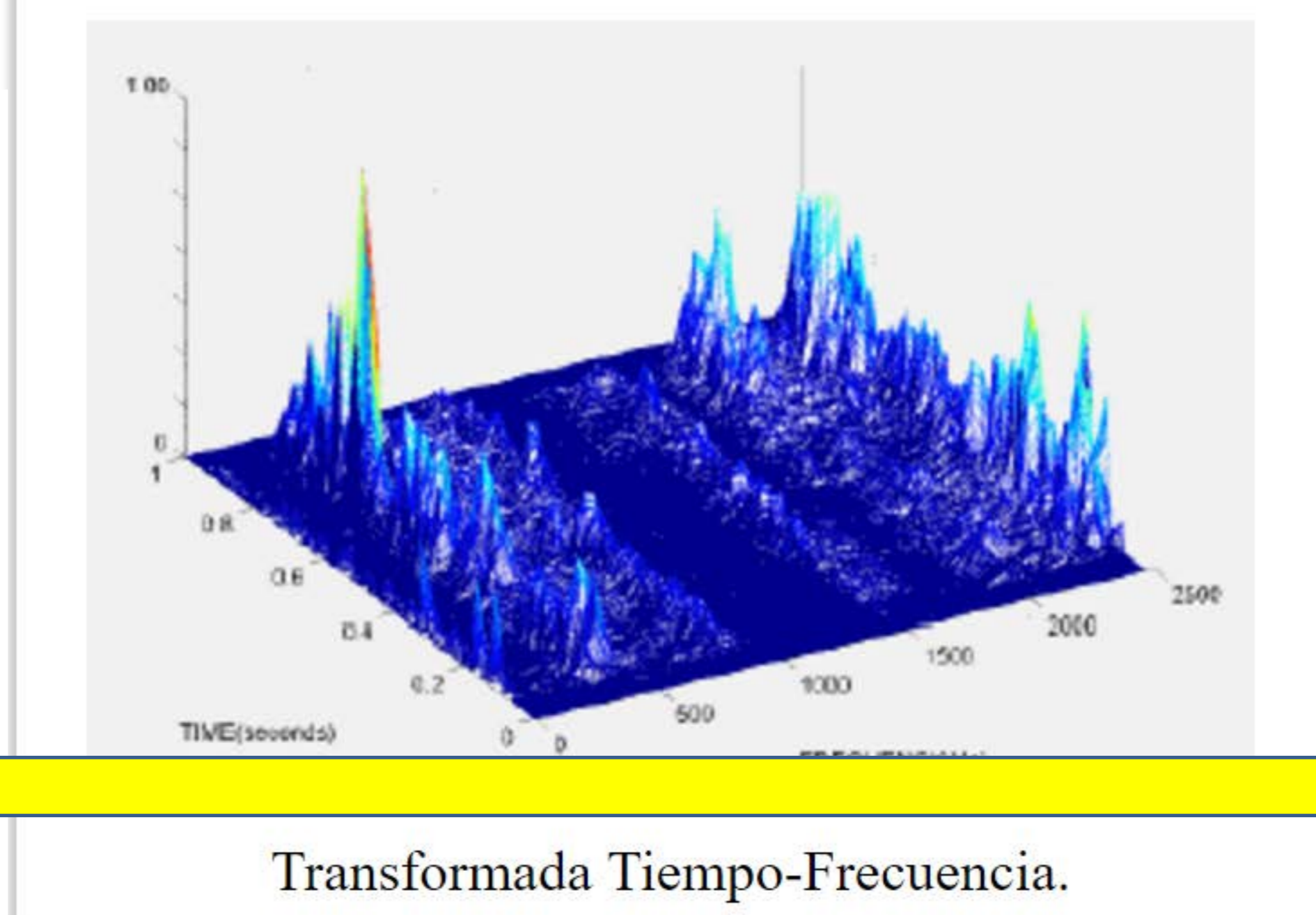
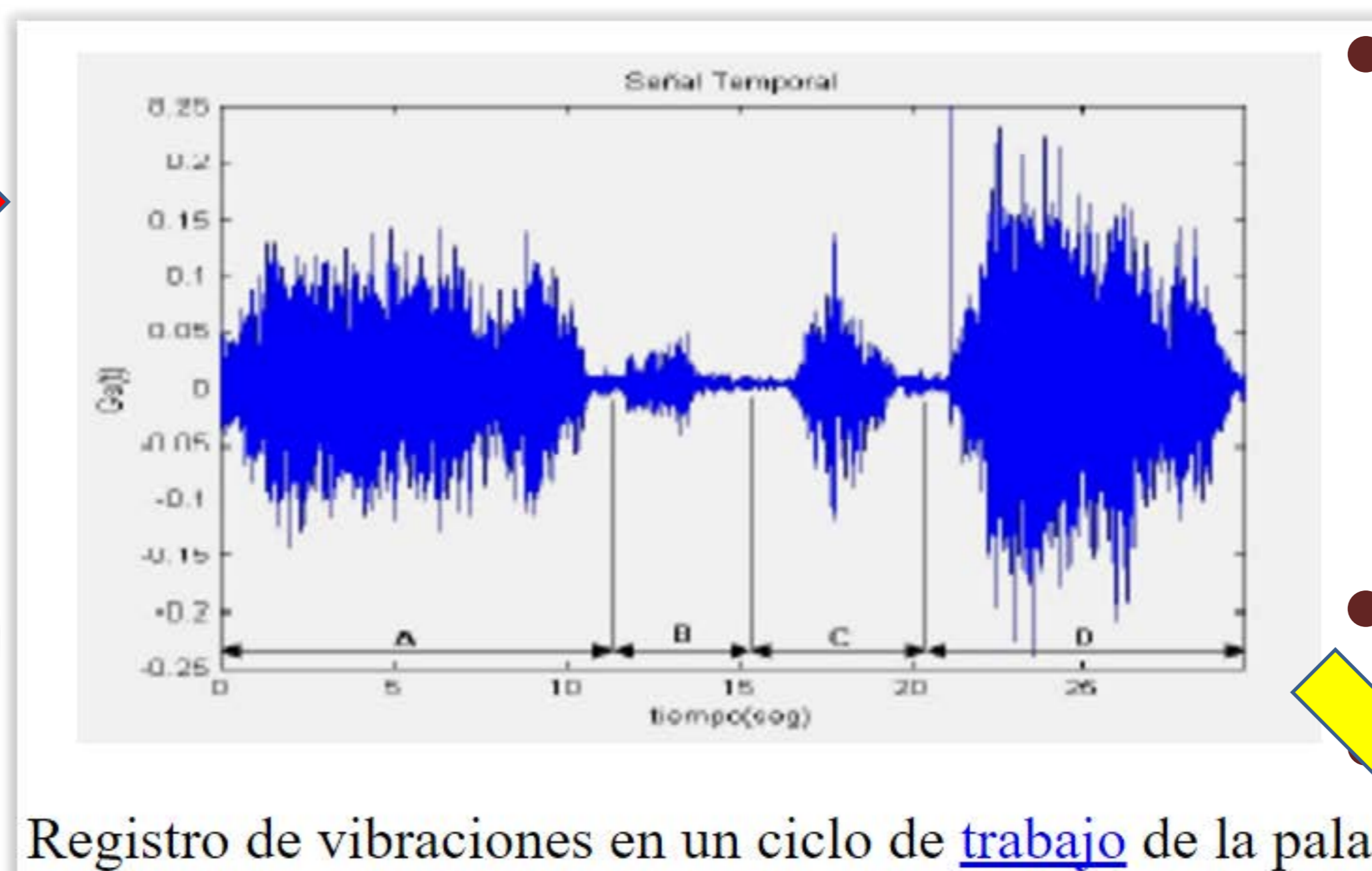
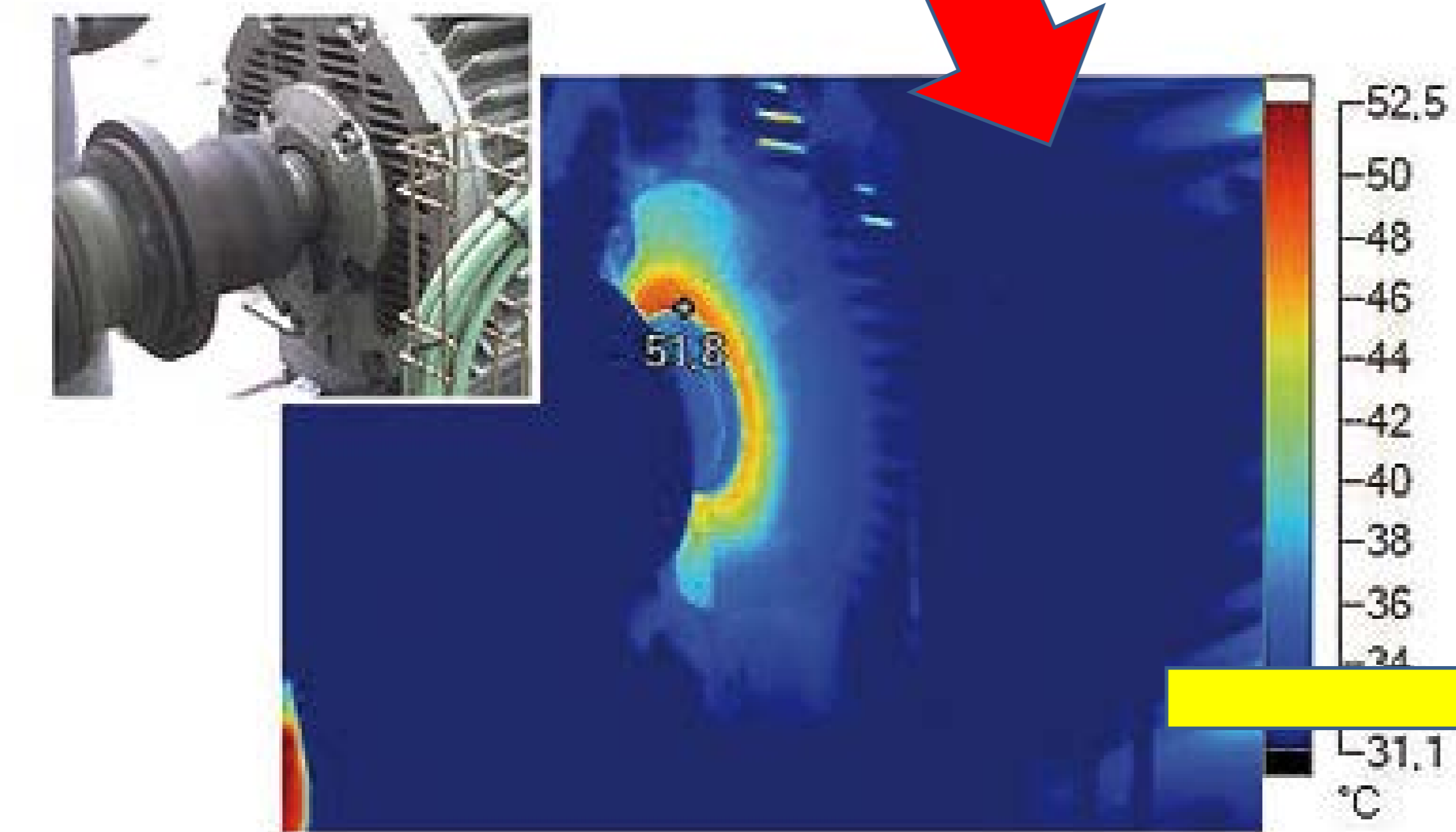


**BOSCH**



# Caso Práctico IV: Mantenimiento Predictivo

- Detectar posibles fallas y defectos de maquinaria en las etapas incipientes para evitar que estos fallos se manifiesten en uno más grande durante su funcionamiento, evitando que ocasionen paros de emergencia y tiempos muertos, causando impacto financiero negativo

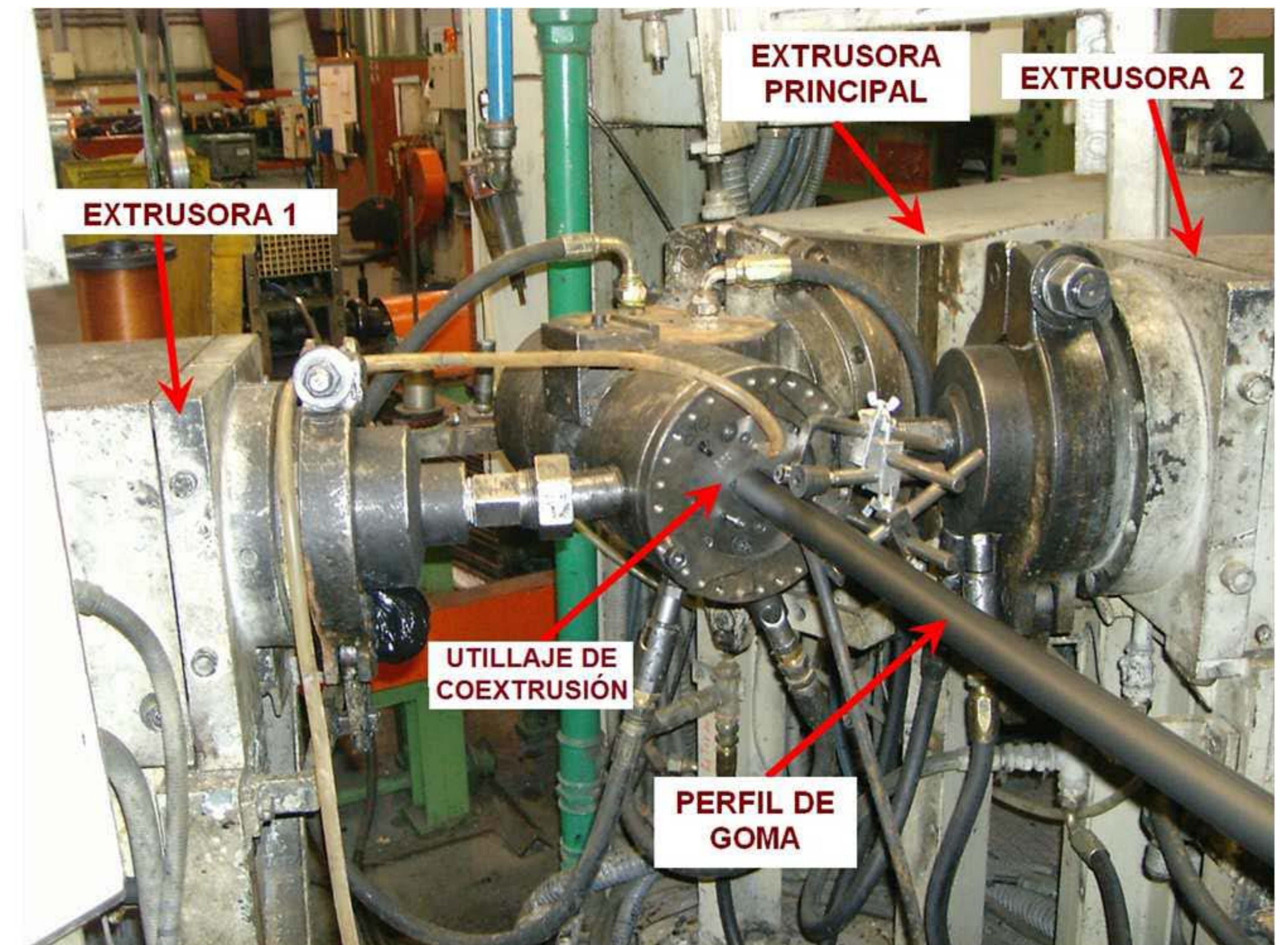
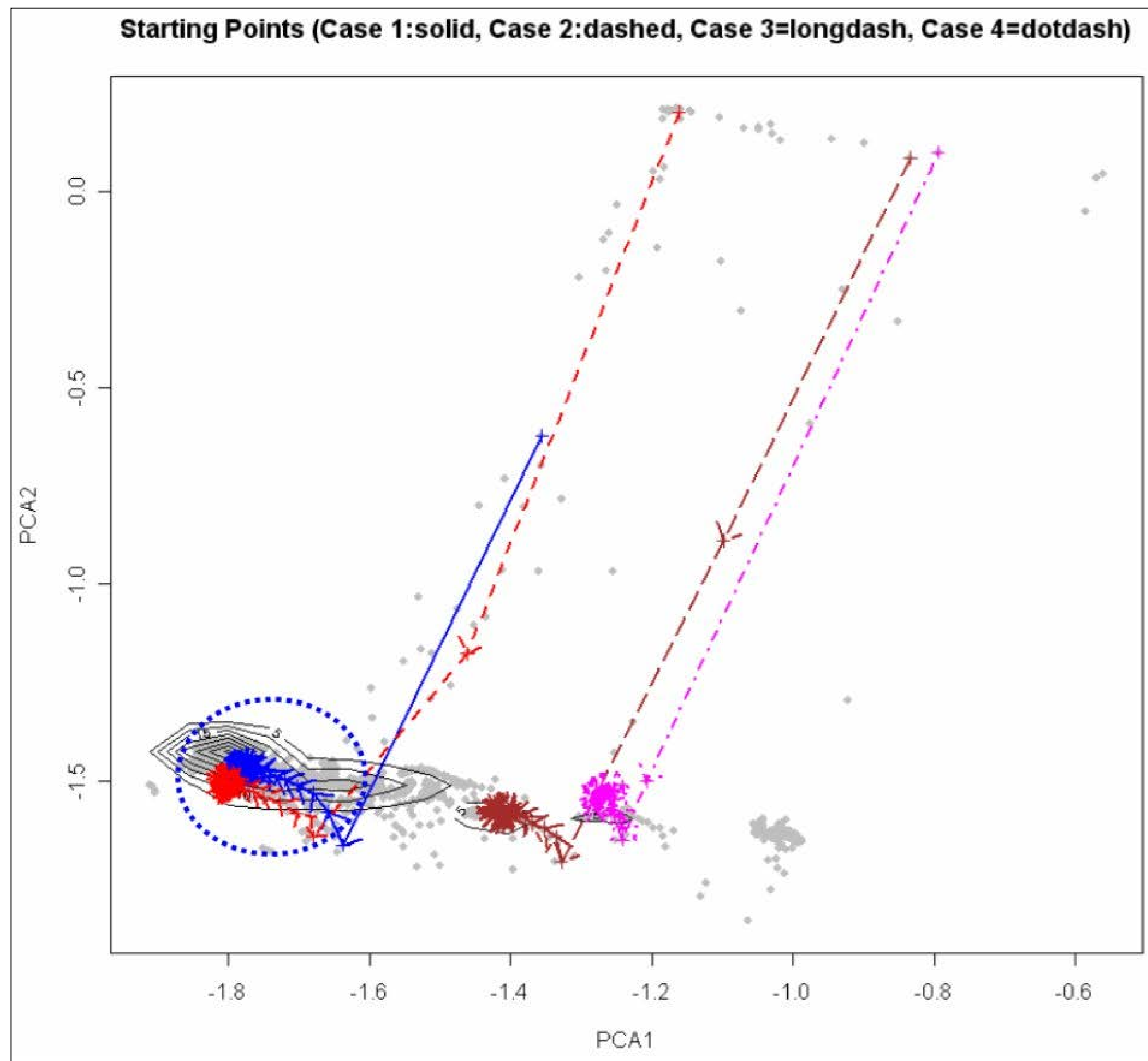


- Monitorización de vibraciones, medida de temperaturas, presiones, etc.
- Cámaras termográficas. Visión artificial.



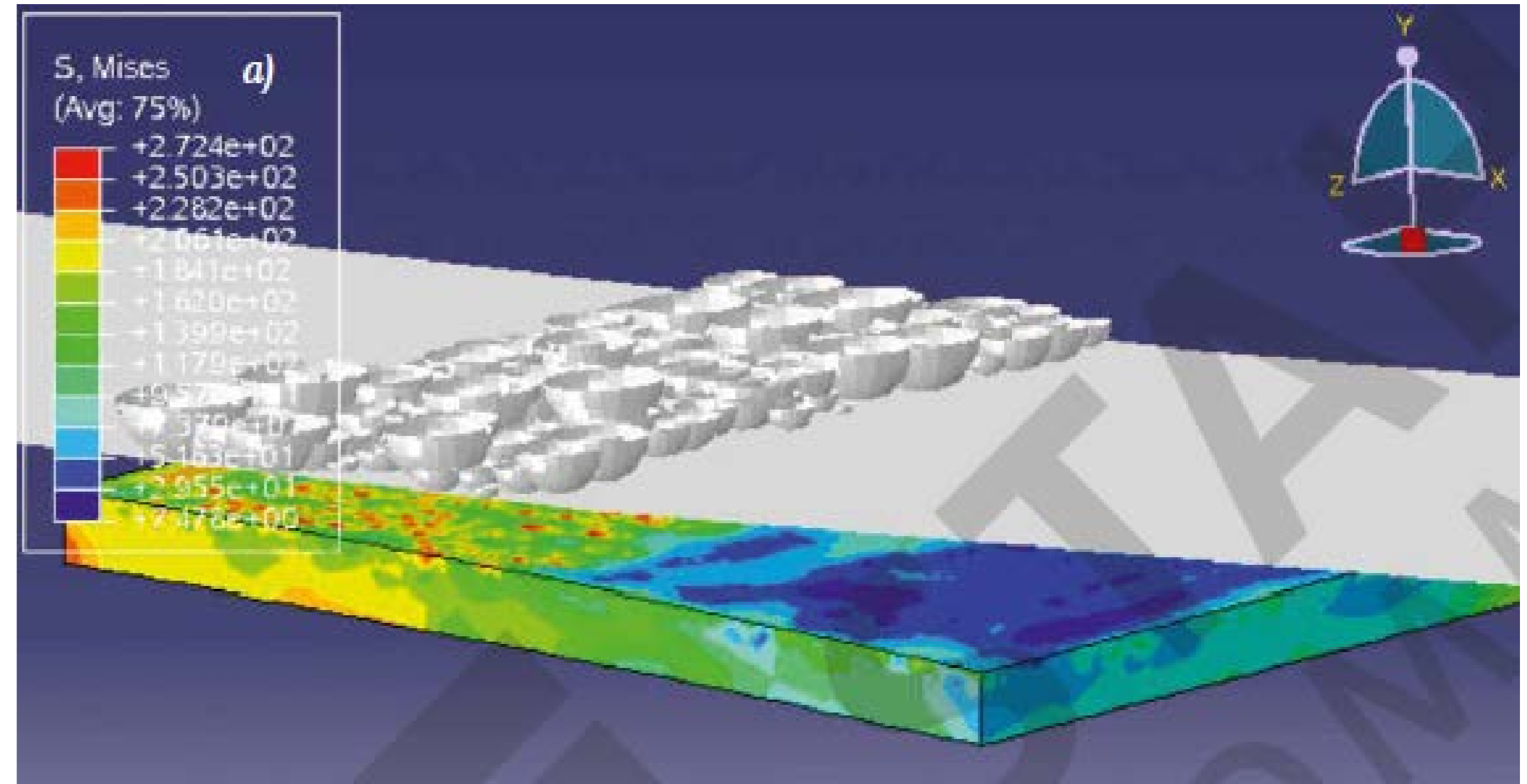
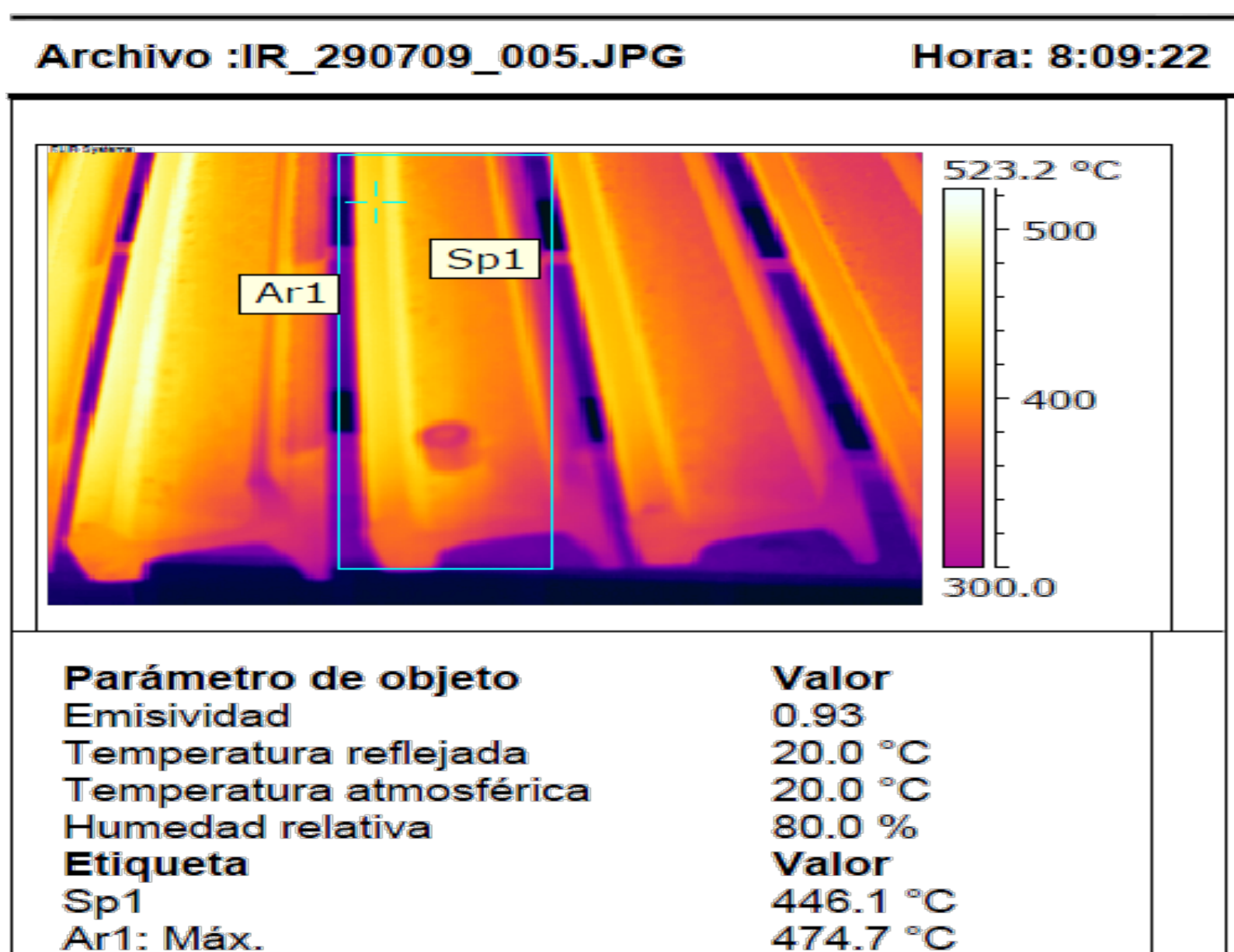
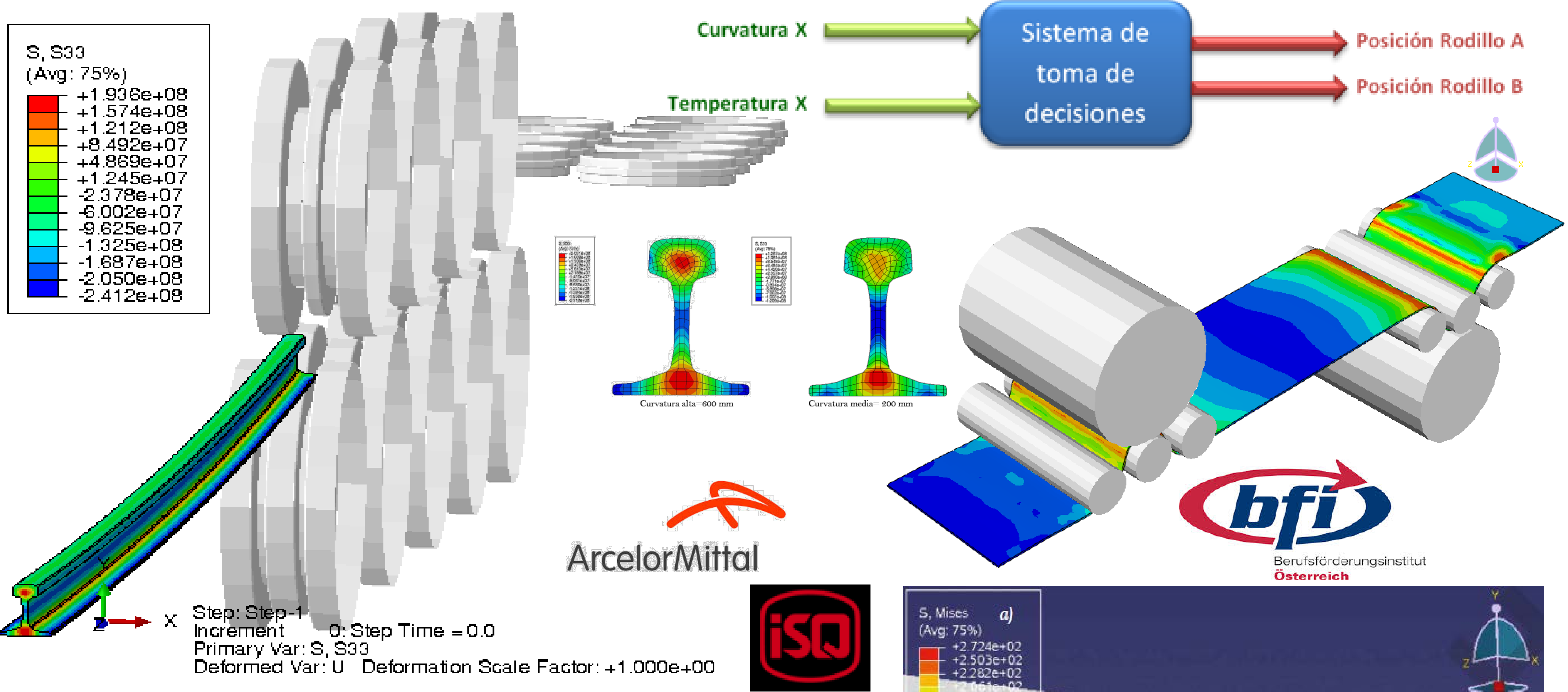
# Caso Práctico V: Monitorización del Punto de Operación

- Monitorización del punto de operación mediante multidimensional scaling (MDS)



DPI2004-07264

# Caso Práctico VI: Combinación Sensórica y FEM/CFD



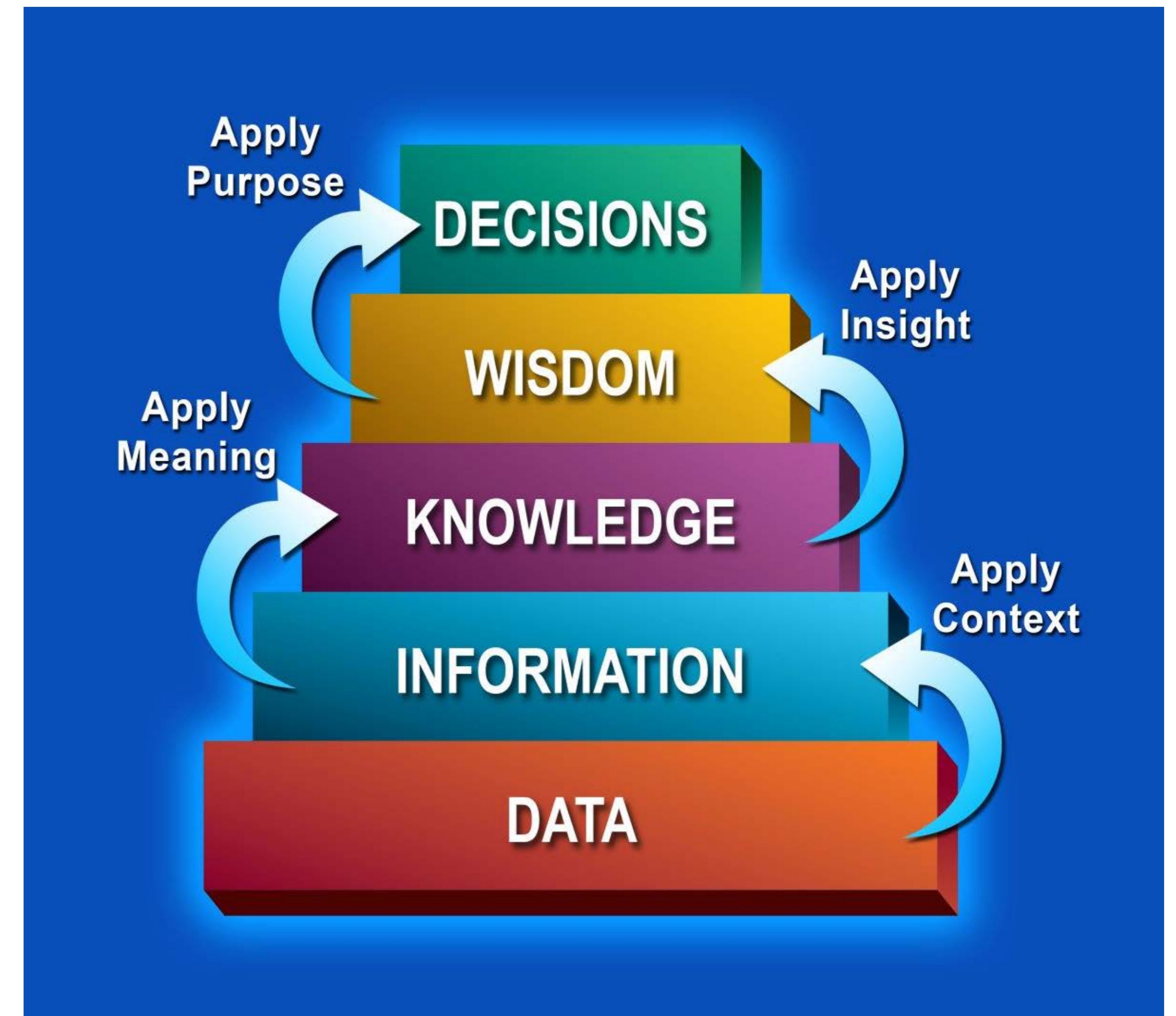
CONTENCA, TESTRA & HYPROCOM National/EU Projects

# ¿Cómo empezar a realizar IBD?

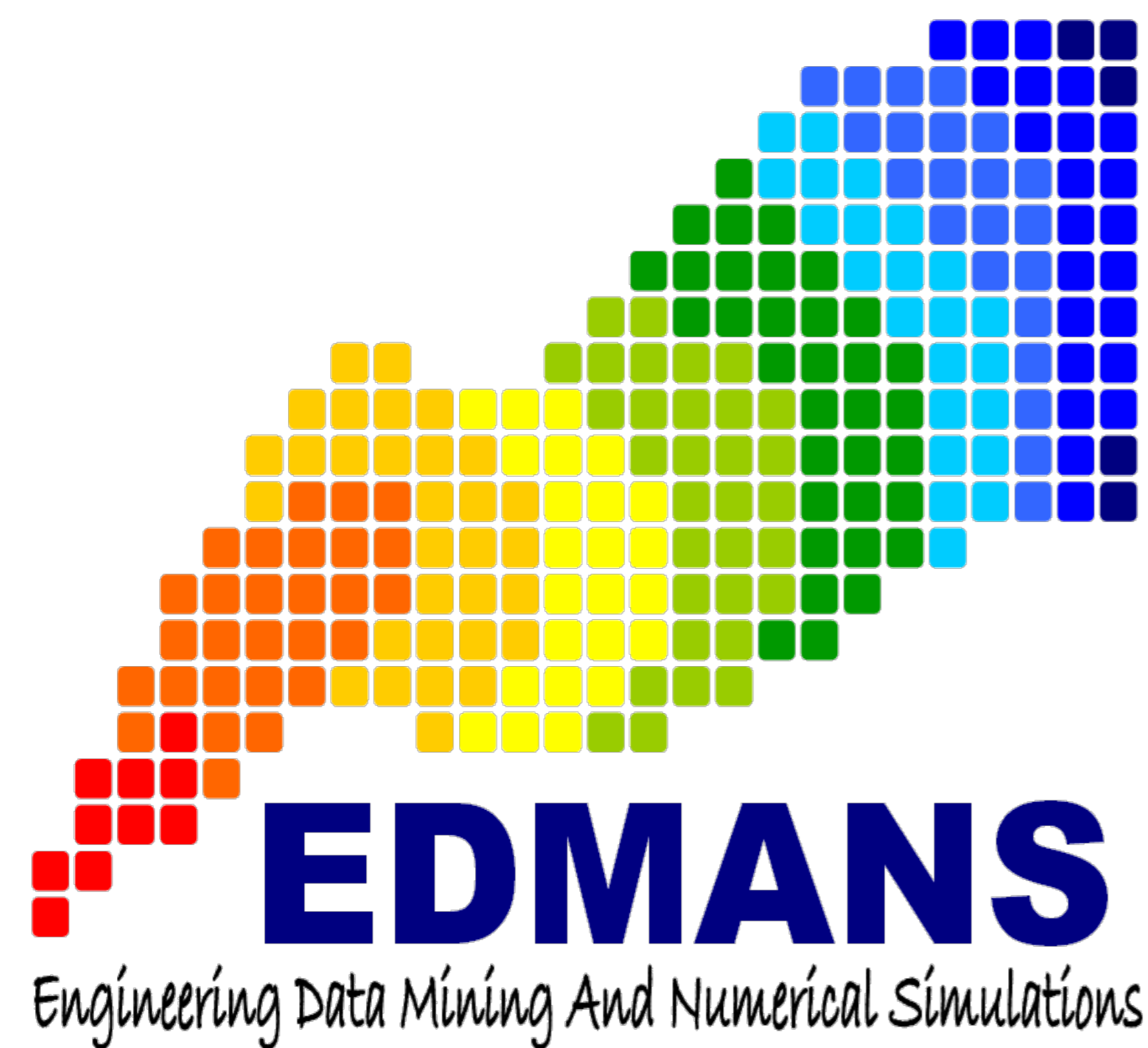
1. ¿Qué podemos hacer? Objetivos. Criterios éxito/fracaso.
2. Análisis de los Datos/Información: ¿Pueden dar valor tal como están?

**Siempre hay que pensar en el VALOR DEL DATO**

1. Diseño y desarrollo del SI para realizar IBD
2. Captura y preprocesado de la Información.
3. **Extracción y validación del Conocimiento.**  
**¿Es útil para la Toma de Decisiones?**
4. Diseño de DSS basado en el conocimiento adquirido
5. Escalar el problema a un problema de IBD en producción
6. Validación y mejora continua del proceso IBD



# ¡MUCHAS GRACIAS!



**UNIVERSIDAD  
DE LA RIOJA**

**Dr. Ing. Ind. Fco. Javier Mtnez de Pisón Ascacibar**

**EDMANS Group**

**fjmartin@unirioja.es**

***<http://mineriadatos.com>***