

Desarrollo de un modelo predictivo para la clusterización y análisis de tendencias en procesos industriales

Javier Podetti, Raizen, javier.podetti@raizen.com.ar

Sinopsis

En el departamento de Integridad de Equipos sometidos a Presión, el análisis de grillas de espesor es una tarea clave para la evaluación de aptitud para el servicio conforme a la norma del American Petroleum Institute (API) 579 “Fitness for Service”. Sin embargo, esta normativa permite únicamente analizar el estado de un componente en un momento específico, sin proporcionar lineamientos para el cálculo de las tasas de corrosión.

Por otro lado, con la incorporación del shale oil en la industria y en la Argentina, resulta fundamental anticipar el comportamiento de nuevos los fenómenos corrosivos asociados.

Este proyecto surge con el objetivo de:

- Determinar rangos de incertidumbre estadísticos a partir de mediciones de espesor sucesivas para un mismo componente.
- Identificar automáticamente patrones en los datos disponibles, con mínima intervención del usuario, facilitando la predicción de la respuesta del sistema ante variaciones operativas, como cambios en la dieta de crudo de una unidad.
- Desarrollar un modelo escalable, independiente del tamaño y tipo de información provista.

Para ello, se trabajó en primera instancia con un conjunto de 10 grillas de mediciones de espesor, cada una con 450 puntos de medición de un recipiente a presión. El análisis de datos mediante algoritmos de clasificación autoexplicativos reduce la dispersión significativa presente en los datos medidos en campo, ayuda a la toma de decisiones, permitiéndonos definir una tasa de corrosión representativa. Junto con el equipo de IT, se desarrolló un prototipo de procesamiento masivo de datos y análisis de tendencias, logrando:

- Reconocer patrones de comportamiento espacial mediante clusterización.
- Determinar de rangos de incertidumbre en la información recopilada.
- Predecir el comportamiento futuro bajo distintos modelos matemáticos (lineal, lineal segmentado, exponencial, cuadrático, menor error, mayor R2).

El objetivo de la herramienta se centró la creación de un prototipo escalable, aplicable a cualquier proceso que requiera la correlación y predicción de parámetros operativos. La versatilidad de este prototipo lo hace adaptable para análisis de datos en diversas áreas, incluyendo:

- Inspección
- Tecnología
- Mantenimiento
- Economía y Planificación

Introducción

En el departamento de Integridad de Equipos sometidos a Presión, el análisis de grillas de espesor es una tarea crítica para la evaluación de aptitud para el servicio conforme a la norma API 579 “Fitness for Service”. Sin embargo, esta normativa permite únicamente analizar el estado de un componente en un momento específico, sin proporcionar lineamientos para el cálculo de las tasas de corrosión.

La incorporación del shale oil en la dieta de crudo de una refinería tiene asociada la incertidumbre de cómo afectará al comportamiento de los mecanismos de degradación activos en las diferentes unidades.

En este contexto, resulta vital contar con una herramienta que permita el análisis veloz de los datos de proceso y a su vez detecte patrones para la predicción del comportamiento a futuro, con mínima intervención del usuario.

El objeto de este trabajo es el desarrollo de un prototipo de procesamiento masivo de datos que incluya un modelo predictivo para la clusterización de comportamientos y análisis de tendencias en procesos industriales, como lo es el análisis de pérdida de espesor en recipientes a presión.

El alcance de este trabajo es detallar el prototipo de procesamiento masivo de datos que incluya un modelo predictivo para la clusterización de comportamientos y análisis de tendencias en procesos industriales. El caso de estudio utilizado fue la corrosión acelerada en el tope de un recipiente acumulador. El objetivo principal de este prototipo es brindarle herramientas de análisis al inspector de la unidad que permitan determinar los planes de inspección y mantenimiento que mantengan la unidad en un nivel de riesgo tan bajo como sea razonablemente practicable (ALARP por sus siglas en inglés). El mismo fue desarrollado en la Refinería Buenos Aires junto con la firma Simpleza SA.

Desarrollo

Inicialmente se contaba con grillas de 450 puntos de medición (15 generatrices divididas en 30 mediciones angulares) de espesores en toda la periferia accesible del cabezal del equipo analizado (ver Figura.1 y Figura.2). Las mediciones consecutivas demostraban una marcada aceleración en la velocidad de corrosión con zonas de corrosión preferencial, producto de un cambio en las variables operativas. Es de particular interés calcular una velocidad de corrosión representativa del componente, ya que los lineamientos bridados por API 579 se centran en el cálculo de espesor de retiro, no de la velocidad de corrosión. En una instancia previa se calculó el espesor de retiro del componente, permitiendo calcular la vida remanente una vez establecida la tasa de corrosión.

VISTA SUPERIOR

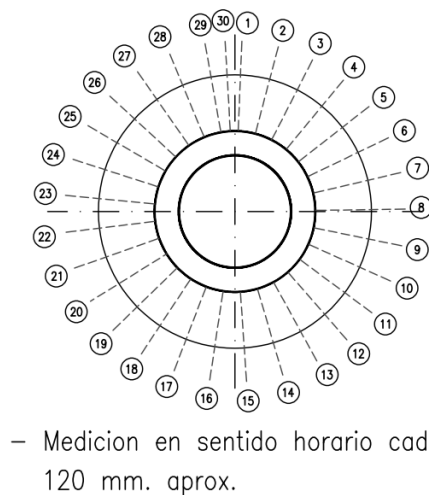


Figura.1: Esquema de medición – Vista superior.

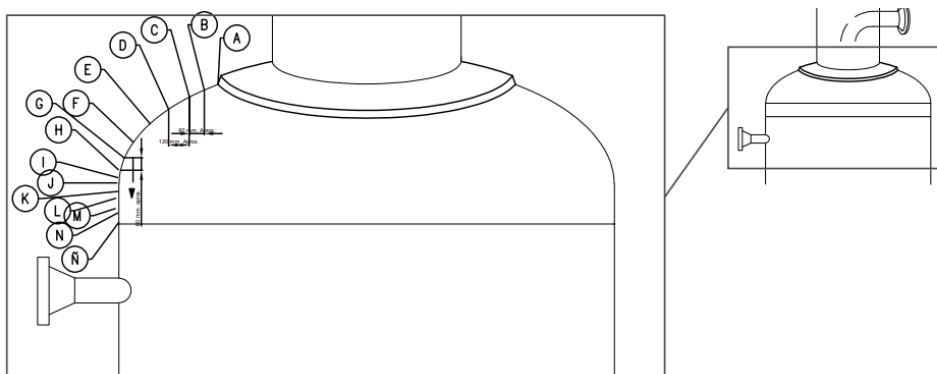


Figura.2: Esquema de medición – Vista en elevación.

Al no contar con instrumentos para la medición continua de espesores en el equipo en cuestión, las grillas se deben realizar mediante palpadores ultrasónicos manuales, con las incertidumbres propias de una técnica manual realizada por un operador. Consecuentemente, previo a la realización de este prototipo, se contaba con un recipiente con evidencia de degradación localizada, siendo inspeccionado por una técnica viable de presentar variabilidad espacial por aspectos como:

- Posición del palpador.
- Equipo utilizado para medir.
- Personal que realiza la medición.

Puntos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
A	12.2	11.9	11.9	12.2	12.0	11.7	11.9	11.6	11.9	12.0	11.7	11.7	12.1	11.5	12.0	11.9	12.1	12.3	10.7	11.2	11.0	10.9	11.1	10.9	11.2	11.2	11.0	11.5	11.2	11.2
B	12.0	12.0	11.6	11.6	11.7	11.7	11.4	11.2	11.3	11.7	12.2	11.8	12.0	12.2	12.0	11.3	12.4	12.2	11.1	11.2	11.1	10.8	11.0	10.9	11.0	11.0	12.2	10.8	11.2	11.5
C	12.3	11.5	11.6	11.6	11.9	11.8	11.2	11.4	11.4	11.9	12.3	11.9	11.8	12.3	11.9	11.8	12.0	11.7	11.2	11.7	10.9	10.8	11.0	11.0	11.3	11.2	12.9	11.4	12.3	11.6
D	12.2	11.9	12.1	11.9	11.8	11.5	11.4	12.1	11.7	11.8	12.5	12.4	12.3	12.5	12.5	11.9	11.9	14.0	11.1	11.6	10.9	10.8	11.2	11.4	11.5	11.5	12.2	11.5	12.4	11.9
E	12.2	12.2	12.3	11.6	11.7	12.0	11.8	12.1	12.3	11.9	12.7	12.9	12.4	12.6	13.1	12.4	11.9	13.9	11.2	11.3	11.2	11.4	11.6	11.2	11.5	11.8	12.0	11.5	12.5	12.0
F	12.5	12.6	12.5	11.8	11.7	11.8	11.3	12.9	12.6	12.0	13.2	13.2	12.5	12.5	13.5	12.9	12.2	12.0	11.4	11.2	11.1	11.0	11.7	11.1	11.8	12.2	12.9	11.4	12.0	12.0
G	12.6	12.6	12.8	12.5	11.8	12.3	11.2	12.5	12.8	12.4	13.5	13.6	12.9	12.9	13.6	13.6	12.7	12.2	11.7	11.0	11.4	11.4	11.7	11.5	12.0	12.2	12.3	11.6	12.4	12.4
H	12.9	13.0	13.2	12.0	11.8	12.5	12.2	12.4	13.0	12.6	13.3	12.8	13.4	12.9	13.0	13.5	12.8	12.9	11.9	11.7	11.7	11.7	12.4	11.6	12.1	12.8	12.5	11.8	12.5	11.9
I	12.8	13.2	13.2	12.2	12.1	12.9	12.2	12.7	13.5	13.1	13.3	12.9	13.0	13.5	13.6	14.0	13.5	12.4	12.2	12.0	11.8	12.4	12.4	12.2	12.4	12.6	12.6	11.4	12.6	12.2
J	13.0	13.4	12.8	12.1	12.1	13.0	12.3	13.3	13.7	13.4	13.4	13.4	13.1	13.0	12.8	13.7	13.1	12.7	11.9	12.2	13.0	12.9	12.4	12.4	13.1	12.8	12.1	12.4	12.0	
K	13.4	13.4	13.2	12.6	12.6	12.5	13.1	13.2	13.3	13.5	12.5	13.3	13.0	13.6	S/A	13.0	13.6	S/A	12.5	12.3	12.4	12.7	13.2	12.4	12.3	12.2	12.5	12.2	12.4	12.1
L	13.6	13.2	13.1	12.6	12.5	13.1	13.0	13.0	13.5	13.8	12.9	12.9	13.3	12.5	S/A	13.1	13.8	S/A	12.7	12.4	12.3	12.9	13.0	12.7	12.8	12.4	12.7	12.5	12.3	12.5
M	13.8	13.6	12.8	13.3	12.3	12.8	13.3	13.4	13.1	13.7	13.6	12.7	13.8	12.6	S/A	13.2	13.3	S/A	12.4	12.5	12.3	12.4	12.5	13.0	13.0	13.3	12.4	12.7	12.5	12.5
N	13.4	13.4	12.8	12.9	12.7	13.2	13.5	13.7	13.5	13.6	13.5	12.4	13.0	12.9	S/A	13.1	13.3	S/A	13.6	12.2	12.5	12.4	13.4	12.8	12.8	13.8	13.9	12.3	12.5	12.4
N	13.3	12.4	12.8	13.0	12.4	13.1	13.2	12.5	13.0	14.0	13.5	12.7	13.0	14.0	S/A	13.0	14.0	S/A	12.7	12.5	12.9	13.3	14.1	12.4	12.4	13.0	13.5	12.2	12.0	12.5

Figura.3: Ejemplo de grilla de espesores.

Conociendo el estado del arte en términos de procesamiento de información y análisis predictivo de tendencias, se contactó a la firma Simpleza SA para desarrollar un prototipo de análisis capaz de retroalimentarse con cada medición consecutiva. El mismo debería ser capaz de determinar los rangos de incertidumbre estadísticos asociados a la predicción, y centrarse en la identificación de superficies con comportamiento similar (clusterización) para analizar su comportamiento mediante medidas como el promedio, la media o la mediana de dichos clusters.

Como resultado de este proyecto, se desarrolló un prototipo de análisis predictivo autoexplicativo¹ que, con mínima intervención del usuario, procesa los conjuntos de datos provistos, genera dos tipos de clusterización independientes, además de predecir el comportamiento de cada punto de medición individual, y alimenta un tablero de PowerBI para visualización de los resultados y toma de decisiones.

En primer lugar, se realizó un análisis exploratorio de la información provista, para descartar anomalías e inconsistencias en las mediciones. Dado que se observaron incrementos en el espesor medido e inconsistencias en las mediciones, se decidió descartar las mediciones donde se observe un incremento mayor a la tolerancia establecida (0.1mm) para reemplazarlas con el valor inmediato anterior. Se agregó una opción para comparar visualmente el efecto de este pre-procesamiento (boton espesor real/ajustado).

El desarrollo del prototipo buscó que el mismo no esté limitado al tipo de información provista ni al tamaño de la matriz de datos. En el futuro, puede ser adaptado para cualquier departamento de la Refinería que busque correlacionar conjuntos de datos de forma automática. El prototipo se desarrolló alrededor de un mecanismo de pérdida de espesor, en el cual la variable observada sólo puede decrecer. En caso de adaptar el prototipo a otros fenómenos a estudiar, este criterio debe ser revisado.

¹ El modelo de análisis busca cuadrar el comportamiento observado con una fórmula matemática dependiente únicamente del tiempo. No se relaciona la variable observada (en este caso, el espesor) con una variable de control (como podría ser la temperatura o el pH) ya que al momento del desarrollo de este prototipo no se contaba con mediciones instrumentadas de otros parámetros operativos de interés.

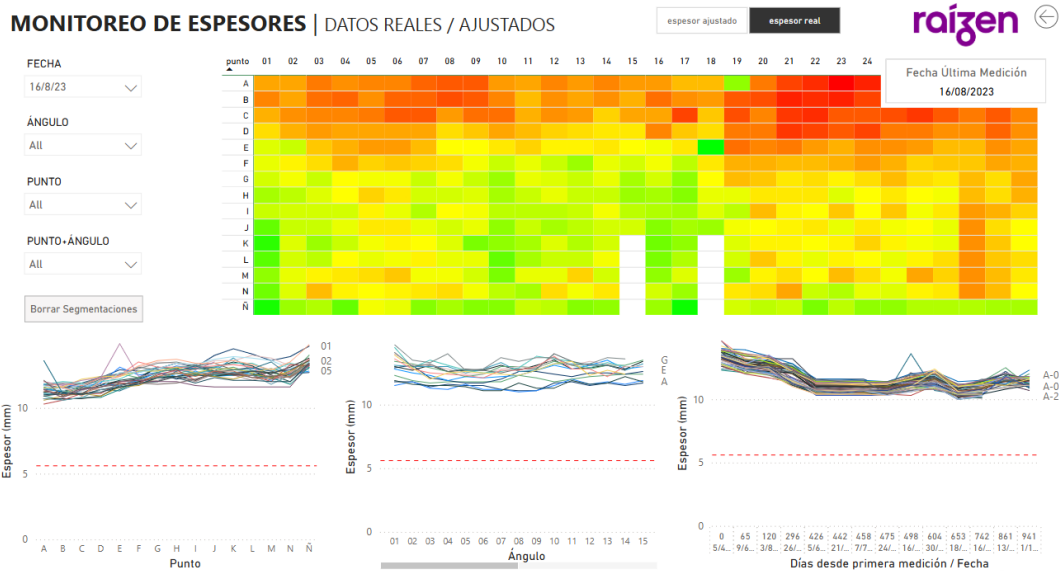


Figura.4 Monitoreo de espesores.

A pedido del Inspector de la unidad, se incluyó una visualización 3D de los espesores medidos con indicaciones de color para simplificar la interpretación de datos.

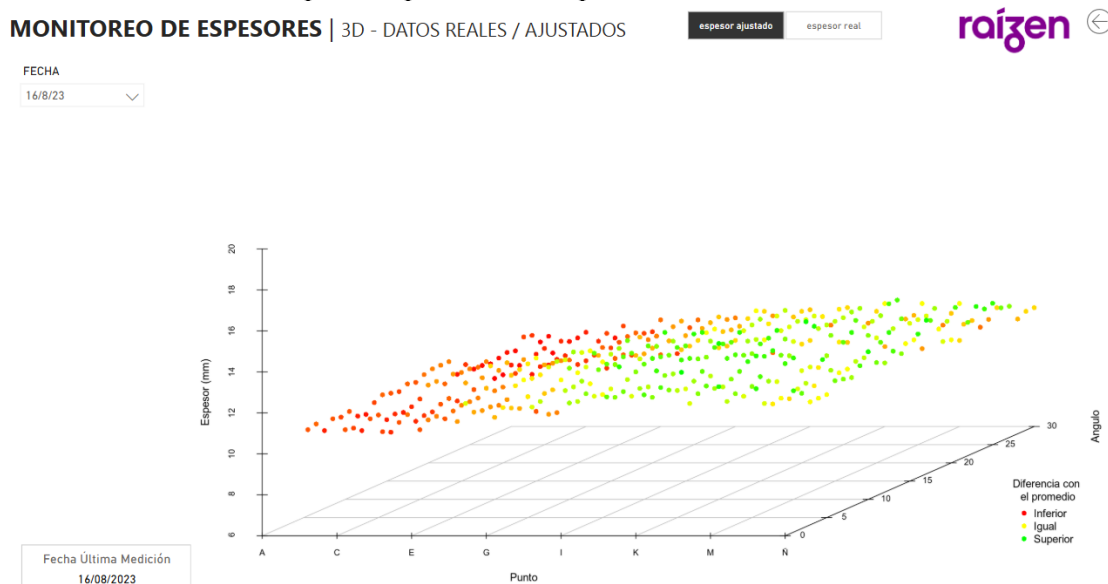


Figura.5 Visualización 3D.

En la Figura.6 se observa el tablero de clusterización automática desarrollado por Simpleza, el cual organiza los puntos de medición en 5 categorías en función de su comportamiento mediante un mapa de colores. Los datos se agruparon utilizando el algoritmo no jerárquico k-means. Se presenta una medida de la proporción de variabilidad de los datos que es explicada por este agrupamiento. También calcula automáticamente la media de los espesores medidos a través del tiempo con los rangos de incertidumbre (+/- dos desviaciones estándar) asociados a la utilización del modelo. Finalmente, el gráfico de barras muestra el tiempo promedio al espesor de retiro (5.6 mm) y el nivel de dispersión dentro de cada grupo (intervalos).

Se observa que:

- Al momento del desarrollo de este prototipo, la pérdida acelerada de espesor había comenzado a estabilizarse.
- La media de los días hasta el umbral se encuentra alrededor de los 3000 días (debido a la estabilización de la pérdida de espesor), pero se cuenta con rangos de incertidumbre tan grandes que el período de tiempo que permitiría estar bajo un nivel de riesgo ALARP se encuentra realmente alrededor de los 900 días.
- En caso de repetirse las condiciones operativas que llevaron a la pérdida acelerada de espesor, no se contaría con el tiempo suficiente para reemplazar el componente.

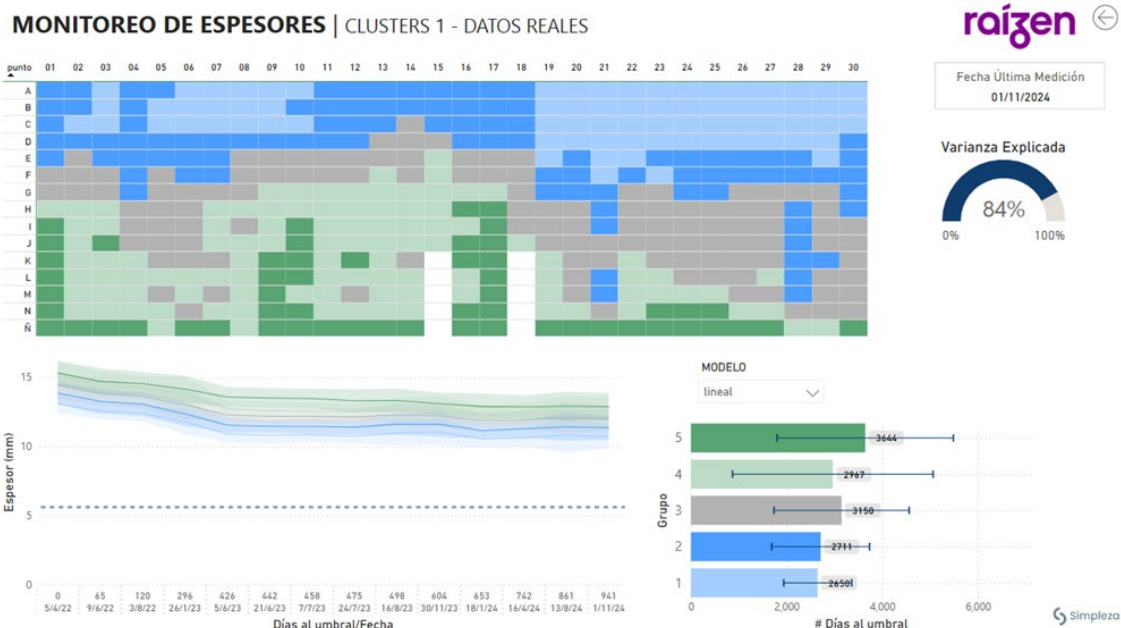


Figura.6 Clusterización automática.

Para aumentar la cantidad de herramientas disponibles para la toma de decisiones por parte del Inspector, se decidió incluir en el prototipo una clusterización “manual”, que divide los comportamientos observados en términos de espesor y velocidad de corrosión en cuartiles. En este agrupamiento se observan medias menores y desviaciones estándar menores, con resultados similares (tiempo al retiro ~ 900 días) que en el modelo de clusterización anterior.

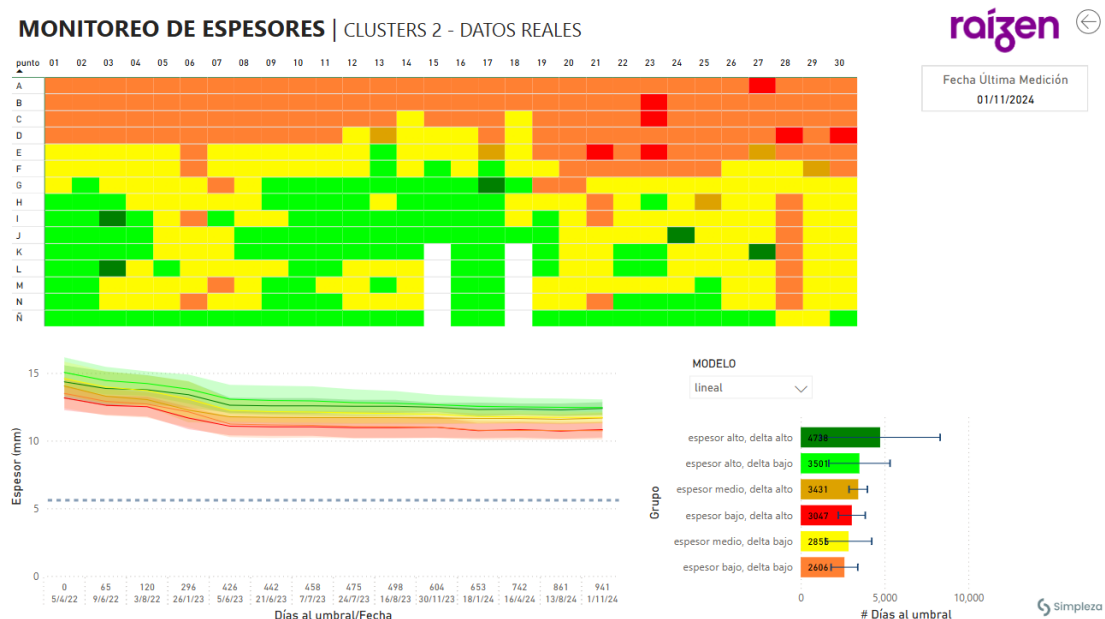


Figura.7 Clusterización manual.

Ya que al analizar la media de un conjunto de datos se pueden enmascarar resultados desfavorables, se desarrolló un tablero que analice el comportamiento de cada punto de medición junto con sus rangos de incertidumbre.

El tablero desarrollado permite variar entre distintos modelos de predicción estadísticos. En este caso se limitó a dos modelos, uno lineal y uno exponencial, ya que la utilización de otros modelos de parametrización no representaba correctamente el comportamiento del componente. Sin embargo, se incorporaron a la programación del prototipo otros modelos matemáticos típicos, como lo son el cuadrático y el lineal segmentado, para futuras aplicaciones. Como

último paso, se incluyó una opción para observar punto a punto el comportamiento del modelo que presente el mayor R^2 o el menor error cuadrático medio.

En la visualización también se incluyó un filtro para analizar el comportamiento de los distintos clusters definidos en los tableros anteriores.

Como es esperable, se observa que los rangos de incertidumbre de las mediciones punto a punto entregan resultados más conservadores que los resultados clusterizados (tiempo al retiro ~ 600 días).

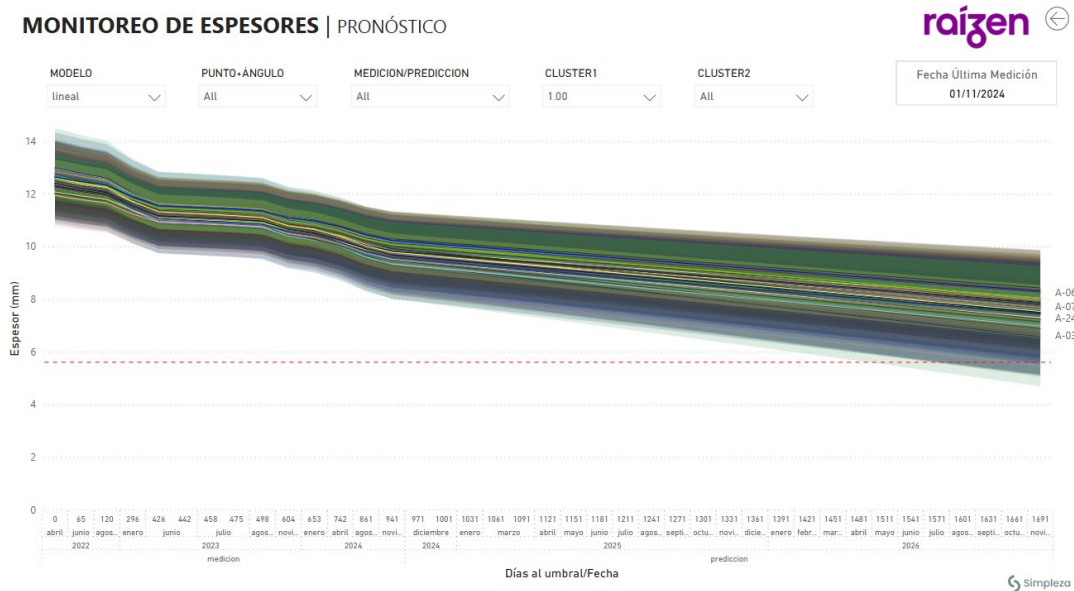


Figura.8 Predicción de comportamiento punto a punto.

Luego de evaluar los resultados obtenidos de los distintos tableros, se confirmó la decisión de reemplazar el equipo en la próxima parada planeada de la unidad. El análisis también permitió aumentar los períodos entre inspecciones, ya que se cuenta con una medida confiable de la velocidad de corrosión y tiempo al retiro y sus intervalos de incertidumbre. La herramienta seguirá siendo válida aún luego de la renovación del componente, y podrá ser utilizada para su monitoreo.

Conclusiones

Se desarrolló un prototipo de procesamiento masivo de datos y análisis de tendencias, logrando:

- Reconocer patrones de comportamiento espacial mediante clusterización.
- Determinar de rangos de incertidumbre en la información recopilada.
- Predecir el comportamiento futuro bajo distintos modelos matemáticos (lineal, lineal segmentado, exponencial, cuadrático, menor error, mayor R^2).

La funcionalidad del prototipo desarrollado puede ser ampliada para diversos campos de aplicación. El desarrollo buscó que el mismo no esté limitado al tipo de información provista ni al tamaño de la matriz de datos. En el futuro, puede ser adaptado para cualquier departamento de la Refinería que busque correlacionar conjuntos de datos de forma automática.

El prototipo se desarrolló alrededor de un mecanismo de pérdida de espesor, en el cual la variable observada sólo puede decrecer. En caso de adaptar el prototipo a otros fenómenos a estudiar, este criterio debe ser revisado.

El modelo desarrollado, como herramienta de soporte en la toma de decisiones, permitió:

- Determinar la vida remanente del equipo estudiado, permitiendo ajustar el plan de inspección y mantenimiento de forma segura y confiable.
- Extender los períodos de inspección, reduciendo los costos.

- Disponer de una herramienta de análisis ágil en caso de modificar las condiciones de operación.

Bibliografía

1. API 579-1/ASME FFS-1 2021 Fitness-For-Service

Breve Cv de los autores

Título y lugar de estudio: Ingeniero Mecánico – Universidad de Buenos Aires

Cargo actual en la empresa: Ingeniero en Recipientes a Presión

Trayectoria (diferentes empresas en las que hizo su carrera):

- **Givaudan S.A (2016 – 2018)**
- **Besna S.A (2019 – 2023)**
- **Raizen (2023 – Actualidad)**