



UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DetECCIÓN DE ANOMALÍAS EN PROCESOS INDUSTRIALES MEDIANTE ANÁLISIS DE CONECTIVIDAD Y MODELOS PREDICTIVOS DE Machine Learning

Aplicación al análisis de sensores en una planta de celulosa

Por: Carolina Elizabeth Valenzuela Carrasco

Tesis presentada a la Facultad de Nombre de Facultad de la Universidad de
Concepción para optar al grado académico de Ingeniero Civil Matemático

Noviembre 2025

Concepción, Chile

Profesor Guía: Guillermo Patricio Ferreira Cabezas

Tutor Guía: Manuel Ignacio Pérez Carrasco

© 2025, Nombre

Ninguna parte de esta tesis puede reproducirse o transmitirse bajo ninguna forma o por ningún medio o procedimiento, sin permiso por escrito del autor.

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento

Agradecimientos

Quiero expresar mi más profundo agradecimiento a mi familia por su apoyo incondicional y por acompañarme en todas mis caídas y victorias. Su confianza y amor fueron el motor que me impulsó a continuar.

A mis profesores, por su disposición a escucharme, por responder cada una de mis dudas y por brindarme un apoyo que trascendió lo académico. Gracias por orientarme, motivarme y contribuir de manera valiosa a mi formación.

A mis compañeros de carrera, gracias por la colaboración, el compañerismo y las incontables horas de estudio compartidas. Nuestros logros son también un reflejo del esfuerzo conjunto.

A mis amigos, gracias por su paciencia, por comprender mis ausencias y por estar siempre presentes sacándome una sonrisa.

A Dios, por darme fortaleza en los momentos difíciles y por guiarme a lo largo de este camino.

Y, finalmente, a quienes ya no están conmigo físicamente, pero siempre permanecerán en mi corazón. Su recuerdo, sus enseñanzas y su amor me acompañaron durante todo este proceso.

A todos ustedes, gracias.

Resumen

Este trabajo aborda el problema de la detección temprana de anomalías en procesos industriales, lo cual es fundamental para prevenir fallas críticas que afectan la continuidad operativa, la eficiencia productiva y los costos operacionales de la empresa. En particular, identificar condiciones anormales antes de que ocurran eventos críticos permite aplicar una estrategia preventiva. Desde el punto de vista técnico, el desafío radica en reconocer patrones que preceden a una falla, los cuales no siempre son evidentes mediante el monitoreo individual de sensores.

En este contexto, se busca anticipar los cortes de hoja en una planta de celulosa mediante el análisis de las relaciones dinámicas entre variables dentro del proceso industrial y la detección de comportamientos anómalos previos al evento, con el objetivo de encontrar la causa raíz de un corte de hoja. La hipótesis central plantea que el uso de técnicas de análisis de conectividad y de modelos autoregresivos permite identificar sensores precursores y comprender la propagación de anomalías en la maquinaria.

El estudio se basa en variaciones industriales del proceso de secado de celulosa de la planta CMPC Santa Fe, monitoreadas por 96 sensores con muestreo cada cinco minutos. En una primera etapa, se aplica el enfoque de conectividad bidireccional de Diebold–Yilmaz mediante modelos VAR-LASSO, con el objetivo de detectar patrones en las relaciones entre sensores y cuantificar su influencia sobre una serie binaria que representa la ocurrencia del corte de hoja. Posteriormente, se implementa un esquema de detección de anomalías utilizando modelos regresivos de Machine Learning, incluyendo K-Nearest Neighbors, Regresión Lineal, Random Forest, Support Vector Regression y XGBoost. Los modelos son entrenados bajo condiciones normales del proceso y evaluados en períodos cercanos al evento, comparando los errores de predicción mediante la métrica MSE.

Los resultados permiten identificar sensores y zonas del proceso con mayor capacidad de detección ante condiciones anómalas previas a los cortes de hoja, aportando información relevante para el monitoreo preventivo. El trabajo demuestra el potencial en el área industrial del uso combinado de análisis de conectividad y aprendizaje automático como herramienta para la detección temprana de fallas y la disminución de sus consecuencias.

Abstract

This work addresses the problem of early anomaly detection in industrial processes, which is essential to prevent critical failures that affect operational continuity, production efficiency, and the company's operational costs. In particular, identifying abnormal conditions before critical events occur enables the implementation of a preventive strategy. From a technical perspective, the challenge lies in recognizing patterns that precede a failure, which are not always evident through the individual monitoring of sensors.

In this context, this study aims to anticipate sheet breaks in a pulp mill by analyzing the dynamic relationships among variables within the industrial process and by detecting anomalous behaviors prior to the event, with the objective of identifying the root cause of a sheet break. The central hypothesis states that the use of connectivity analysis techniques and autoregressive models allows the identification of precursor sensors and the understanding of anomaly propagation throughout the machinery.

The study is based on industrial variations from the pulp drying process at the CMPC Santa Fe plant, monitored by 96 sensors with a five-minute sampling interval. In the first stage, the bidirectional connectivity approach of Diebold–Yilmaz is applied using VAR-LASSO models, with the objective of detecting patterns in sensor relationships and quantifying their influence on a binary series representing the occurrence of a sheet break. Subsequently, an anomaly detection scheme is implemented using regression-based Machine Learning models, including K-Nearest Neighbors, Linear Regression, Random Forest, Support Vector Regression, and XGBoost. The models are trained under normal process conditions and evaluated during periods close to the event, with prediction errors compared using the MSE metric.

The results allow the identification of sensors and process areas with greater detection capability under anomalous conditions preceding sheet breaks, providing relevant information for preventive monitoring. This work demonstrates the potential of the combined use of connectivity analysis and machine learning in the industrial domain as a tool for early failure detection and the reduction of its consequences.

Índice general

Agradecimientos	I
Resumen	II
Abstract	III
1 Introducción	1
1.1 Contexto	1
1.2 Planteamiento del problema	2
1.2.1 Cortes de hoja	3
1.2.1.1 Causas principales del Corte de Hoja	3
1.3 Objetivos	4
1.3.1 Objetivo general	4
1.3.2 Objetivos específicos	4
2 Marco Teórico	5
2.1 Series de Tiempo y Estacionariedad	5
2.1.1 Series de Tiempo Multivariada	5
2.1.2 Estacionariedad	6
2.1.2.1 Test Dickey-Fuller Aumentado	7
2.1.2.2 Aplicación en R	7
2.1.3 Modelos de Vectores Autorregresivos de orden 1	8
2.1.4 Modelos de Vectores Autorregresivos de orden p	10
2.1.5 LASSO-VAR	13
2.2 Análisis de Conectividad con Spillovers	15
2.2.1 Concepto Spillovers	15
2.2.2 Diebold–Yilmaz	15
2.2.2.1 Matriz de Conectividad	17
2.3 Modelos de Regresión para Detección de Anomalías	18
2.3.1 Correlación de Pearson	19
2.3.2 Regresión Lineal (LR)	19
2.3.3 Support Vector Regression (SVR)	20
2.3.3.1 Hiperparámetros clave	20
2.3.4 Random Forest (RF)	20
2.3.4.1 Hiperparámetros clave	20
2.3.5 K-Nearest Neighbors (KNN)	21

2.3.5.1	Hiperparámetros clave	21
2.3.6	XGBoost	21
2.3.6.1	Hiperparámetros clave	21
2.4	Validación de modelos	22
2.4.1	Validación cruzada	22
2.4.1.1	Cálculo del error	23
2.4.2	Validación Hold-out	23
2.4.2.1	Cálculo del error	24
3	Metodología	25
3.1	Datos	25
3.1.1	Fuente de Datos	25
3.1.2	Sensores	26
3.1.3	Tiempos perdidos	31
3.1.4	Bitácora Cortes de hoja	32
3.1.5	Serie Binaria CH	35
3.2	Análisis de conectividad	35
3.2.1	Metodología Spillovers	35
3.2.2	Preprocesamiento	36
3.2.2.1	Tratamiento de series constantes	36
3.2.2.2	Estacionariedad	37
3.2.3	Ranking de sensores con mayor conectividad hacia CH	37
3.3	Entrenamiento con datos normales	38
3.3.1	Construcción de Máscaras Booleanas	38
3.3.2	Preprocesamiento	39
3.3.3	Entrenamiento con time series split	40
3.3.4	Cálculo MSE	41
3.3.5	Selección de hiperparámetros	42
3.3.6	Definición de umbral de detección	43
3.4	Entrenamiento con datos reales	44
3.4.1	Validación hold out	44
3.4.2	Cálculo del MSE por ventana	46
3.4.3	Monitoreo temporal del MSE y activación de umbrales de alerta	47
3.4.4	Indicador global de desempeño	47
3.4.5	Selección del modelo óptimo	48
4	Resultados	49
4.1	Conectividad	49
4.1.1	Ranking	49
4.1.2	Correlación	51
4.2	Desempeño de modelos de entrenamiento y selección de hiperparámetros	53
4.2.1	Umbrales de detección	53
4.2.2	Conclusiones preliminares de las tablas de detección de umbrales	59

4.3	Identificación del modelo óptimo	60
5	Conclusión	64
5.1	Conclusión	64
	Referencias	66
	Apéndices	68
A	Resultados extras	68
A1	Monitoreo temporal MSE	69
A2	Máscaras booleanas de entrenamiento y test	83

Índice de cuadros

2.2.1 Esquema de matriz de conectividad	17
3.1.1 Descripción sensores	29
3.1.2 Descripción sensores	30
3.1.3 Cortes de hoja 2025	34
3.3.1 Orden de prioridad por cercanía al origen de la máquina	40
3.3.2 Posibles parámetros	42
3.4.1 Ventanas train y test para el entrenamiento	46
4.1.1 Ranking de sensores según su conectividad hacia la variable de corte de hoja (CH). Para cada columna CH_k , la columna n° representa la posición del ranking, donde $n^\circ = 1$ identifica al sensor con mayor conectividad dirigida hacia CH_k . En cada celda se indica el número del sensor, y entre paréntesis el valor de la conectividad estimada.	51
4.1.2 Correlación Pearson >0.9	52
4.2.1 Sensores que superaron el umbral $\times 5$ para el modelo KNN . () es el n° de ventanas en que fue superado el umbral $\times 5$	54
4.2.2 Sensores que superaron el umbral $\times 5$ para el modelo LR . () es el n° de ventanas en que fue superado el umbral $\times 5$	55
4.2.3 Sensores que superaron el umbral $\times 5$ para el modelo RF . () es el n° de ventanas en que fue superado el umbral $\times 5$	56
4.2.4 Sensores que superaron el umbral $\times 5$ para el modelo SVR . () es el n° de ventanas en que fue superado el umbral $\times 5$	57
4.2.5 Sensores que superaron el umbral $\times 5$ para el modelo XGB . () es el n° de ventanas en que fue superado el umbral $\times 5$	58
4.3.1 Proporción de sensores que superaron el umbral $F = 5$ para cada modelo y evento de corte de hoja.	61
4.3.2 Valor global g_m de cada modelo.	62

Índice de figuras

3.1.1	Máquina secadora FLAKT y distribución de las etapas del proceso.	26
3.1.2	Sensor 75 de Tensión y 96 de Velocidad cercano al Corte de Hoja 1.	32
3.1.3	Serie de tiempo de los cortes de hoja 2025	35
3.3.1	Máscara operacional para el año 2025	39
3.3.2	Mapa de etapas de la máquina secadora	39
3.3.3	Error de reconstrucción (MSE) del modelo XGBoost con diferentes combinaciones de parámetros, evaluado por fold	43
4.2.1	Evolución temporal del error del modelo KNN para el CH-2, indicando los niveles de umbral superados.	59
4.3.1	Evolución temporal del error del modelo LR para el CH-3, indicando los niveles de umbral superados.	62
A1.1	Comparación métodos CH-1	69
A1.2	Comparación vertical de métodos en CH-2	70
A1.3	Comparación vertical de métodos en CH-3	71
A1.4	Comparación vertical de métodos en CH-4	72
A1.5	Comparación vertical de métodos en CH-5	73
A1.6	Comparación vertical de métodos en CH-6	74
A1.7	Comparación vertical de métodos en CH-7	75
A1.8	Comparación vertical de métodos en CH-8	76
A1.9	Comparación vertical de métodos en CH-9	77
A1.10	Comparación vertical de métodos en CH-10	78
A1.11	Comparación vertical de métodos en CH-11	79
A1.12	Comparación vertical de métodos en CH-12	80
A1.13	Comparación vertical de métodos en CH-13	81
A1.14	Comparación vertical de métodos en CH-14	82
A2.1	Máscaras booleanas CH-1	83
A2.2	Máscaras booleanas CH-2	83
A2.3	Máscaras booleanas CH-3	84
A2.4	Máscaras booleanas CH-4	84
A2.5	Máscaras booleanas CH-5	85
A2.6	Máscaras booleanas CH-6	85
A2.7	Máscaras booleanas CH-7	86
A2.8	Máscaras booleanas CH-8	86
A2.9	Máscaras booleanas CH-9	87
A2.10	Máscaras booleanas CH-10	87

A2.11 Máscaras booleanas CH-11	88
A2.12 Máscaras booleanas CH-12	89
A2.13 Máscaras booleanas CH-13	89
A2.14 Máscaras booleanas CH-14	90

Capítulo 1

Introducción

1.1. Contexto

La **celulosa** es la **materia prima** para una amplia variedad de productos derivados del papel, como papel higiénico, cartón, servilletas y empaques sustentables, siendo uno de los **pilares de las exportaciones forestales de Chile**, [CMPC \(2025\)](#). Chile, por su parte, mantiene una posición destacada en el comercio internacional de pulpa. En 2023, las exportaciones chilenas de celulosa alcanzaron los US\$2.690 millones, lo que representó un **crecimiento interanual del 18,9 %** [de Relaciones Económicas Internacionales \(2024\)](#).

Esto consolidó a la **pulpa química** como el **producto más relevante** dentro de las exportaciones forestales chilenas. Solo en el primer trimestre de 2025, las exportaciones de pulpa química totalizaron US\$904,9 millones, representando un **56,5 % del total exportado** por el sector [Forestal \(2025\)](#).

1.2. Planteamiento del problema

Sin embargo, este proceso enfrenta desafíos técnicos importantes, uno de los eventos críticos es el **corte de hoja**, una anomalía que consiste en la **ruptura o separación repentina de la hoja continua de celulosa**. Esta anomalía provoca la **detención inmediata de la operación**, generando pérdidas de materia prima, aumento en los tiempos de inactividad y una disminución considerable en la eficiencia de la planta, lo que interrumpe el proceso de secado y genera detenciones de la línea continua de la máquina y pérdidas económicas ([ciencia y tecnología \(2021\)](#))).

La **detección temprana de patrones que preceden a estos eventos**, mediante el análisis de datos generados por **sensores industriales** y el uso de modelos predictivos basados en inteligencia artificial, representa una herramienta estratégica para avanzar hacia una industria más eficiente ([Emma \(2025\)](#))).

El presente proyecto propone el desarrollo de un sistema de análisis de conectividad dinámica entre variables críticas, con el objetivo de identificar causas precursoras de los cortes de hoja utilizando técnicas estadísticas multivariadas.

Esta aproximación no solo busca mejorar la capacidad de reacción operativa frente a posibles fallas, sino que también contribuye al conocimiento en torno al monitoreo inteligente, el mantenimiento predictivo y la detección de anomalías en el sistema de secado de una planta de celulosa.

1.2.1. Cortes de hoja

La fabricación de celulosa involucra una cadena de procesos industriales complejos y continuos. Una de las etapas finales es el proceso de secado, donde la hoja de celulosa es transportada por una máquina secadora de gran escala. Sin embargo, este proceso puede verse interrumpido por una anomalía crítica conocida como corte de hoja, que consiste en el rompimiento abrupto de la hoja continua de celulosa.

El corte de hoja implica la detención inmediata de la línea de producción. Dado que se trata de un **proceso en cadena**, este evento genera un impacto significativo, se detiene la operación completa de los diversos procesos, se incrementan los tiempos de inactividad, se pierde materia prima, y se produce una disminución en la eficiencia operacional, lo que genera grandes pérdidas económicas.

1.2.1.1. Causas principales del Corte de Hoja

Entre las principales causas que pueden originar un corte de hoja se han identificado las siguientes:

- Variaciones en la velocidad de la máquina secadora.
- Cambios abruptos en la humedad o presión del sistema.
- Problemas mecánicos, como el desgaste o desalineación de los rodillos.
- Errores en la dosificación o concentración de productos químicos.
- Fallas humanas.

Este proyecto se focaliza en el análisis de la etapa de secado mediante sus sensores, se busca comprender la dinámica interna del sistema desde una perspectiva de series temporales multivariadas, buscando detectar patrones de comportamiento que preceden a los cortes de hoja.

1.3. Objetivos

1.3.1. Objetivo general

Analizar la **conectividad entre los sensores** del proceso de secado y la **ocurrencia de cortes de hoja** para identificar los sensores más influyentes como **posibles precursores** de estos eventos, y aplicar modelos de regresión supervisada a estos sensores para predecir su comportamiento normal detectando desviaciones cercanas a los eventos anómalos y así **anticipar la ocurrencia de cortes** en tiempo real.

1.3.2. Objetivos específicos

- Realizar un **análisis exploratorio de los sensores** y su comportamiento.
- **Cuantificar el grado de conectividad entre sensores** mediante la metodología de Spillovers con el enfoque de Diebold-Yilmaz.
- Identificar sensores que actúan como **emisores fuertes de influencia (shocks)** hacia la serie de corte de hoja.
- Comparar las conectividades entre distintos cortes de hoja para **detectar patrones recurrentes**.
- **Elaborar un ranking de los sensores más influyentes** sobre la ocurrencia de los cortes de hoja.
- Aplicar modelos de regresión supervisada (KNN, SVR, Random Forest y regresión lineal) a cada sensor del ranking para **predecir su comportamiento bajo condiciones normales**, evaluar desviaciones y detectar posibles anomalías cercanas al corte de hoja.
- Generar un **umbral máximo de error en condiciones normales** para evaluar el comportamiento del error cercano al corte de hoja y sus desviaciones de este umbral.

Capítulo 2

Marco Teórico

2.1. Series de Tiempo y Estacionariedad

2.1.1. Series de Tiempo Multivariada

Las series de tiempo multivariadas van más allá del análisis de series univariadas al considerar **múltiples variables observadas simultáneamente** a lo largo del tiempo. Esto permite capturar tanto la dependencia temporal dentro de cada serie como las **interdependencias entre ellas**, lo que resulta esencial en sistemas complejos donde **las variables pueden influenciarse mutuamente**.

Un reciente estudio destaca que el análisis multivariado puede aprovechar patrones y correlaciones entre variables, generando modelos de predicción más precisos y robustos que los univariados, [Kim et al. \(2025\)](#). Además, el uso de múltiples variables permite que el modelo incorpore información relevante que influye en la serie principal, mejorando así la capacidad predictiva.

El enfoque Series de Tiempo Multivariadas es relevante en aplicaciones reales que implican múltiples factores interrelacionados como precios internacionales, métricas financieras, condiciones ambientales o eventos económicos disruptivos, [McCowan \(2024\)](#).

Los principales beneficios del enfoque multivariado son:

- **Mayor precisión predictiva:** se aprovechan las correlaciones entre series para capturar mejor la dinámica conjunta.
- **Causalidad:** permite estudiar cómo unas variables pueden conducir a otras en el tiempo, así como analizar efectos de dependencia.
- **Adaptabilidad:** el modelado multivariado es más robusto frente a perturbaciones externas, eventos macroeconómicos o choques inesperados, ya que incorpora variables exógenas relevantes.

2.1.2. Estacionariedad

Una serie de tiempo es estacionaria cuando sus **propiedades estadísticas permanecen constantes a lo largo del tiempo**. En particular, una serie es estacionaria cuando su media y varianza son constantes, y si su covarianza depende únicamente del desfase o rezago entre observaciones, y no del instante temporal en que se mida. Gráficamente, esto implica que los valores de la serie fluctúan alrededor de una media constante y que su variabilidad no cambia sistemáticamente en el tiempo ([Hamilton \(1994\)](#)).

Por el contrario, una serie no estacionaria presenta cambios en su tendencia y/o en su variabilidad. Esto puede manifestarse como un crecimiento o decrecimiento persistente, o como cambios en la amplitud de las fluctuaciones, lo que impide que la serie oscile alrededor de un valor constante ([Box et al. \(2015\)](#)).

Formalmente, una serie de tiempo $\{X_t\}$ es débilmente estacionaria si cumple las siguientes condiciones:

$$E(X_t) = E(X_{t+k}) = \mu$$

$$V(X_t) = V(X_{t+k}) = \sigma^2$$

$$\gamma_k = E[(X_t - \mu)(X_{t+k} - \mu)]$$

donde $E(X_t)$ es la media, $V(X_t)$ la varianza y $E[(X_t - \mu)(X_{t+k} - \mu)]$ la covarianza γ_k entre X_t y X_{t+k} , separados por k periodos.

2.1.2.1. Test Dickey-Fuller Aumentado

Para evaluar la estacionariedad de una serie de tiempo se utiliza el test de Dickey-Fuller Aumentado, propuesto originalmente por [Dickey and Fuller \(1979\)](#) y extendido posteriormente por [Said and Dickey \(1984\)](#).

La estacionariedad es una propiedad fundamental en el análisis de series temporales, ya que **muchas técnicas de modelamiento requieren que los datos no presenten tendencias ni varianzas cambiantes a lo largo del tiempo.**

El test ADF permite contrastar la existencia de raíces unitarias, es decir, determinar si una serie es no estacionaria. La prueba se basa en el siguiente modelo autorregresivo de primer orden:

$$y_t = \phi y_{t-1} + \epsilon_t$$

Si se resta a ambos lados de la ecuación y_{t-1} obtenemos que

$$\Delta y_t = (\phi - 1)y_{t-1} + \epsilon_t \quad (2.1.1)$$

La ecuación 2.1.1 es la base de la prueba Dickey Fuller-Aumentada. Si contamos con una raíz unitaria $\phi = 1$ y el parámetro $\gamma = (\phi - 1) = 0$ se puede aplicar mínimos cuadrados ordinarios al modelo en las primeras inferencias y sin constante. El coeficiente estimado del término autorregresivo se puede someter a la prueba de significancia estadística t para contrastar la hipótesis nula de raíz unitaria contra la alternativa de estacionariedad:

$$H_0 : \gamma = 0 \iff \phi = 1$$

$$H_1 : \gamma < 0 \iff \phi < 1$$

Donde la hipótesis nula es que la serie no es estacionaria y la hipótesis alternativa lo contrario.

2.1.2.2. Aplicación en R

En el software R, el test ADF puede aplicarse mediante la función `adf.test()` del paquete `tseries` ([Trapletti and Hornik \(2023\)](#)). La implementación básica es la siguiente:

```
library(tseries)
adf.test(serie)
```

donde `serie` corresponde a la serie temporal que se desea evaluar. La función retorna, entre otros elementos, el valor p asociado a la prueba. Si este valor p es menor que un nivel de significancia, se rechaza la hipótesis nula de raíz unitaria y se concluye que la serie es estacionaria. En caso contrario, no es posible descartar la no estacionariedad.

2.1.3. Modelos de Vectores Autorregresivos de orden 1

Los modelos de vectores autorregresivos (VAR) permiten analizar la dinámica conjunta de múltiples series de tiempo cuando existe interdependencia entre ellas. Este enfoque es especialmente útil cuando no es posible asumir que algunas variables sean estrictamente externas, ya que el VAR trata todas las series de forma simétrica, permitiendo capturar efectos contemporáneos y rezagados entre las variables (Lütkepohl (2005)).

Consideremos un proceso VAR de primer orden, denotado VAR(1), compuesto por dos variables y_t y z_t . Su representación estructural puede escribirse como:

$$y_t = b_{10} + b_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + e_{y_t}, \quad (2.1.2)$$

$$z_t = b_{20} + b_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + e_{z_t}. \quad (2.1.3)$$

Se asume que y_t y z_t son procesos estacionarios, y que los errores e_{y_t} y e_{z_t} son ruidos blancos independientes, con varianzas σ_y^2 y σ_z^2 , respectivamente. En este sistema ambas variables son internas, puesto que dependen la una de la otra (Hamilton (1994)).

La forma matricial del modelo estructural de las ecuaciones 2.1.2 y 2.1.3 se expresa como:

$$BX_t = \Gamma_0 + \Gamma_1 X_{t-1} + e_t, \quad (2.1.4)$$

donde,

$$B = \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}, \quad X_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}, \quad \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}, \quad \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}.$$

Debido a la presencia de términos contemporáneos entre y_t y z_t , el sistema estructural no es identificable directamente. Para obtener la forma reducida del VAR(1), se multiplica 2.1.4 por la inversa de B :

$$X_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 X_{t-1} + B^{-1}e_t.$$

Si definimos $\Phi_0 = B^{-1}\Gamma_0$, $\Phi_1 = B^{-1}\Gamma_1$ y $u_t = B^{-1}e_t$, obtenemos la forma reducida del VAR(1):

$$X_t = \Phi_0 + \Phi_1 X_{t-1} + u_t, \quad (2.1.5)$$

lo que equivale a:

$$y_t = \phi_{10} + \phi_{11}y_{t-1} + \phi_{12}z_{t-1} + u_{1t},$$

$$z_t = \phi_{20} + \phi_{21}y_{t-1} + \phi_{22}z_{t-1} + u_{2t}.$$

A diferencia del modelo estructural, los errores de la forma reducida (u_{1t}, u_{2t}) se encuentran correlacionados.

A continuación desarrollamos la expresión definida anteriormente:

$$\begin{aligned} \Phi_0 &= B^{-1}\Gamma_0 \\ \begin{pmatrix} \phi_{10} \\ \phi_{20} \end{pmatrix} &= \begin{pmatrix} \frac{1}{1-b_{21}b_{12}} & \frac{b_{12}}{1-b_{21}b_{12}} \\ \frac{b_{21}}{1-b_{21}b_{12}} & \frac{1}{1-b_{21}b_{12}} \end{pmatrix} \begin{pmatrix} b_{10} \\ b_{20} \end{pmatrix} = \begin{pmatrix} \frac{b_{10}+b_{12}b_{20}}{1-b_{12}b_{21}} \\ \frac{b_{20}+b_{21}b_{10}}{1-b_{12}b_{21}} \end{pmatrix} \end{aligned} \quad (2.1.6)$$

De la misma forma, desarrollamos la igualdad:

$$\Phi_1 = B^{-1}\Gamma_1$$

$$\begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{1-b_{21}b_{12}} & \frac{b_{12}}{1-b_{21}b_{12}} \\ \frac{b_{21}}{1-b_{21}b_{12}} & \frac{1}{1-b_{21}b_{12}} \end{pmatrix} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} \frac{\gamma_{11}+b_{12}\gamma_{21}}{1-b_{12}b_{21}} & \frac{\gamma_{12}+b_{12}\gamma_{22}}{1-b_{12}b_{21}} \\ \frac{\gamma_{21}+b_{21}\gamma_{11}}{1-b_{12}b_{21}} & \frac{\gamma_{22}+b_{21}\gamma_{12}}{1-b_{12}b_{21}} \end{pmatrix} \quad (2.1.7)$$

Finalmente,

$$u_t = B^{-1}e_t$$

$$\begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} = \begin{pmatrix} \frac{1}{1-b_{21}b_{12}} & \frac{b_{12}}{1-b_{21}b_{12}} \\ \frac{b_{21}}{1-b_{21}b_{12}} & \frac{1}{1-b_{21}b_{12}} \end{pmatrix} \begin{pmatrix} e_{yt} \\ e_{zt} \end{pmatrix} = \begin{pmatrix} \frac{e_{yt}+b_{12}e_{zt}}{1-b_{12}b_{21}} \\ \frac{e_{zt}+b_{21}e_{yt}}{1-b_{12}b_{21}} \end{pmatrix} \quad (2.1.8)$$

De 2.1.6, 2.1.7 y 2.1.8, obtenemos:

$$\phi_{10} = \frac{b_{10} + b_{12}b_{20}}{1 - b_{12}b_{21}}, \phi_{20} = \frac{b_{20} + b_{21}b_{10}}{1 - b_{12}b_{21}}$$

$$\phi_{11} = \frac{\gamma_{11} + b_{12}\gamma_{21}}{1 - b_{12}b_{21}}, \phi_{12} = \frac{\gamma_{12} + b_{12}\gamma_{22}}{1 - b_{12}b_{21}}, \phi_{21} = \frac{\gamma_{21} + b_{21}\gamma_{11}}{1 - b_{12}b_{21}}, \phi_{22} = \frac{\gamma_{22} + b_{21}\gamma_{12}}{1 - b_{12}b_{21}}$$

$$e_{1t} = \frac{e_{yt} + b_{12}e_{zt}}{1 - b_{12}b_{21}}, e_{2t} = \frac{e_{zt} + b_{21}e_{yt}}{1 - b_{12}b_{21}}$$

Es conocido que $E(\epsilon_{1t}) = E(\epsilon_{2t}) = 0$ y covarianza entre los errores está dada por:

$$E[\epsilon_{1t}, \epsilon_{2t}] = E \left[\left(\frac{e_{yt} + b_{12}e_{zt}}{1 - b_{12}b_{21}} \right) \left(\frac{e_{zt} + b_{21}e_{yt}}{1 - b_{12}b_{21}} \right) \right] = E \left[\frac{e_{yt}e_{zt} + b_{21}b_{21}e_{yt}e_{zt}^2 + b_{21}e_{yt}^2}{(1 - b_{12}b_{21})^2} \right]$$

$$= E \left[\frac{b_{21}\sigma_y^2 + b_{12}\sigma_z^2}{(1 - b_{12}b_{21})^2} \right] \neq 0.$$

2.1.4. Modelos de Vectores Autorregresivos de orden p

Los modelos de Vectores Autorregresivos (VAR) permiten analizar la dinámica conjunta de múltiples series de tiempo cuando existe interdependencia entre ellas. A diferencia de modelos univariados o estructurales donde algunas variables se tratan como externas, en un VAR todas las variables son consideradas internas.

Un VAR resulta especialmente útil en contextos donde no es posible determinar a priori una estructura causal estricta entre las variables, ya que su formulación asume un tratamiento simétrico entre ellas. Esto facilita estudiar cómo cada serie responde tanto a sus propios rezagos como a los rezagos del resto de las series involucradas.

Consideremos un proceso VAR de primer orden, VAR(1), compuesto por dos

variables Y_t y Z_t , se puede expresar como:

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + \epsilon_t \quad (2.1.9)$$

donde Φ_0 es un vector de constantes de dimensión $n \times 1$, Φ_j es una matriz de coeficientes autorregresivos de tamaño $n \times n$ para cada rezago $j = 1, 2, \dots, p$, y ϵ_t corresponde a un vector de ruidos blancos.

Recordemos además que estos ruidos blancos cumplen:

$$E(\epsilon_t) = 0 \quad (2.1.10)$$

$$E(\epsilon_t \epsilon'_\lambda) = \begin{cases} \Omega, & \text{si } t = \lambda \\ 0, & \text{si } t \neq \lambda \end{cases} \quad (2.1.11)$$

Donde Ω es una matriz simétrica y definida positiva de dimensión $n \times n$. A partir del proceso VAR(p), la primera fila de la ecuación del sistema puede expresarse como:

$$\begin{aligned} y_{1t} = & \phi_{10} + \phi_{11}^{(1)} y_{1,t-1} + \phi_{12}^{(1)} y_{2,t-1} + \cdots + \phi_{1n}^{(1)} y_{n,t-1} \\ & + \phi_{11}^{(2)} y_{1,t-2} + \phi_{12}^{(2)} y_{2,t-2} + \cdots + \phi_{1n}^{(2)} y_{n,t-2} \\ & + \phi_{11}^{(p)} y_{1,t-p} + \phi_{12}^{(p)} y_{2,t-p} + \cdots + \phi_{1n}^{(p)} y_{n,t-p} \end{aligned}$$

Se observa que un vector autorregresivo corresponde a un sistema de ecuaciones en el cual cada variable es regresionada sobre una constante y sobre sus propios p rezagos, así como sobre los p rezagos de las demás variables del VAR. Esto se aprecia, ya que todas las ecuaciones del sistema comparten el mismo conjunto de variables explicativas. Por otra parte, el modelo VAR(p) presentado en 2.1.9 puede reescribirse utilizando el operador de rezago:

$$Y_t = \left[\frac{\Phi_0}{I_n - \Phi_1 L - \Phi_2 L^2 - \cdots - \Phi_p L^p} \right] + \left[\frac{1}{I_n - \Phi_1 L - \Phi_2 L^2 - \cdots - \Phi_p L^p} \right] \epsilon_t \quad (2.1.12)$$

donde $\Phi(L) = I_n - \Phi_1 L - \Phi_2 L^2 - \cdots - \Phi_p L^p$ es una matriz de $n \times n$ que se agrupa a los n polinomios característicos del modelo VAR(p). A partir de la expresión 2.1.12, se observa que:

$$E[Y_t] = \left[\frac{\Phi_0}{I_n - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p} \right] = \mu$$

Con todo lo anterior, el modelo VAR(p) expresado en desvíos respecto de la media puede escribirse como:

$$(Y_t - \mu) = \Phi_1(Y_{t-1} - \mu) + \Phi_2(Y_{t-2} - \mu) + \dots + \Phi_p(Y_{t-p} - \mu) + \epsilon$$

Por esta razón, la matriz de varianzas y autocovarianzas del proceso queda dada por:

$$\Gamma_j = E[(Y_t - \mu)(Y_{t-j} - \mu)]$$

En un proceso univariante $AR(p)$ se cumple que $\gamma_j = \gamma_{-j}$, reflejando una estructura simétrica en las autocovarianzas. Sin embargo, en un modelo multivariante VAR(p) esta propiedad ya no se mantiene: la reacción de y_{1t} ante una perturbación en y_{2t} no coincide necesariamente con la respuesta inversa. En contraste con el caso univariante, donde para un proceso AR(p) se cumple que $\gamma_j = \gamma_{-j}$, en un modelo VAR(p) esta simetría ya no necesariamente se mantiene. En particular, las covarianzas dirigidas pueden diferir, lo que se refleja en que $\Gamma_j \neq \Gamma_{-j}$.

Bajo este marco, la matriz de covarianzas de un proceso VAR(p) se define como:

$$\Gamma'_j = \Gamma_{-j},$$

donde:

$$\Gamma_j = E[(Y_{t+j} - \mu)(Y_t - \mu)'], \quad \Gamma'_j = E[(Y_t - \mu)(Y_{t+j} - \mu)'] = \Gamma_{-j}.$$

Un proceso $VAR(p)$ será estacionario si el polinomio característico asociado cumple:

$$\det(I_n - \Phi_1 L - \Phi_2 L^2 - \dots - \Phi_p L^p) = 0$$

con raíces tales que $|L| > 1$, es decir, están ubicadas fuera del círculo unitario. Equivalentemente, la condición de estacionariedad puede expresarse mediante:

$$\det(I_n \lambda^p - \Phi_1 \lambda^{p-1} - \Phi_2 \lambda^{p-2} - \dots - \Phi_p) = 0,$$

donde el proceso será estacionario si todas las raíces satisfacen $|\lambda| < 1$. Como

señalan [Demirer et al. \(2018\)](#), cuando el número de variables es elevado, es decir posee alta dimensionalidad, la estimación de un VAR completo se vuelve inviable. En modelos multivariantes suelen existir numerosas variables poco significativas, aunque es difícil identificarlas a priori. Esto ocurre porque el espacio paramétrico crece cuadráticamente con el número de series. De hecho, para un VAR con K variables y p rezagos, el número total de parámetros a estimar es $K + pK^2$ (equivalentemente, $1 + pK$ por cada ecuación), lo que complica la estimación y el pronóstico en sistemas de gran tamaño.

2.1.5. LASSO-VAR

El modelo LASSO-VAR corresponde a una extensión del modelo de vectores autorregresivos (VAR) aplicada a series de tiempo multivariadas de alta dimensión. Este enfoque resulta adecuado cuando se trabaja con un gran número de variables, ya que contiene penalizaciones que reducen el impacto de coeficientes poco significativos.

Un VAR se denomina *Sparse VAR* cuando solo un subconjunto reducido de variables son significativas sobre el sistema. Una de las penalizaciones más utilizadas es la penalización tipo LASSO. Este método tiende a funcionar mejor cuando el conjunto de variables incluye componentes que pueden ser irrelevantes o ser poco significativas.

La regresión LASSO fuerza a algunos coeficientes a ser cero, eliminando así variables no significativas y simplificando el modelo. Por ello, el *LASSO* puede considerarse también como un método de selección automática de variables ([Ferrari et al. \(2021\)](#)). Además, la naturaleza *Sparse* de este enfoque facilita la interpretación de las relaciones entre variables, pudiendo representarse mediante grafos que permiten visualizar las relaciones entre covariables del sistema.

Consideremos un modelo Sparse VAR estacionario en covarianza de N variables para el vector x_t :

$$x_t = \sum_{i=1}^p \Phi_i x_{t-i} + \epsilon_t \quad (2.1.13)$$

Donde $\epsilon_t \sim (0, \Sigma)$ es un vector de perturbaciones no autocorrelacionadas. La representación de media móvil para x_t es:

$$x_t = \sum_{i=0}^{\infty} A_i \epsilon_{t-i} \quad (2.1.14)$$

Donde las matrices A_i de coeficientes $N \times N$ obedecen la recursión $A_i = \Phi_1 A_{i-1} + \Phi_2 A_{i-2} + \dots + \Phi_p A_{i-p}$ con A_0 siendo la matriz identidad y $A_i = 0$ para $i < 0$. Los coeficientes de media móvil son relevantes para entender la dinámica del sistema y las estadísticas de conectividad.

A continuación, considera el estimador de mínimos cuadrados,:

$$\hat{\Phi} = \arg_{\Phi} \min \sum_{t=1}^T (x_t - \sum_{i=1}^p \Phi_i x_{t-i})^2 \quad (2.1.15)$$

sujeto a la restricción: $\sum_{i=1}^p |\Phi_i|^q \leq c$

De manera equivalente, considere el problema de estimación penalizada:

$$\hat{\Phi} = \arg_{\Phi} \min \left(\sum_{t=1}^T (y_t - \sum_i \Phi_i x_{it})^2 + \lambda \sum_{i=1}^K |\Phi_i|^q \right) \quad (2.1.16)$$

Donde LASSO resuelve la regresión penalizada $q = 1$. Las penalizaciones suavemente convexas (por ejemplo, $q = 2$) producen reducciones similares a la regresión Ridge.

El LASSO resuelve el problema de regresión penalizada con L1:

$$\hat{\Phi}_{Lasso} = \arg_{\Phi} \min \left(\sum_{t=1}^T (y_t - \sum_i \Phi_i x_{it})^2 + \lambda \sum_{i=1}^K |\Phi_i| \right) \quad (2.1.17)$$

Lasso realiza reducción y selección de variables. Para ello, utiliza el valor más pequeño de q que garantiza que el problema de minimización sea convexo, lo que facilita los cálculos computacionales

2.2. Análisis de Conectividad con Spillovers

2.2.1. Concepto Spillovers

El concepto de spillover describe **cómo una perturbación en una variable puede influir directa o indirectamente, en otras variables dentro de un sistema** dinámico a lo largo del tiempo, generando efectos en cadena.

Una forma didáctica de comprender este fenómeno es a través de la analogía con un ecosistema:

En un hábitat natural hay diferentes especies como conejos y zorros que interactúan entre sí. Si disminuye la población de conejos (presa principal de los zorros), inevitablemente se verá afectada la población de los zorros. Este tipo de dinámica refleja cómo una alteración en una parte del sistema puede tener consecuencias sobre otros elementos del mismo sistema.

En el contexto industrial de una planta de celulosa, los distintos sensores que monitorean variables como velocidad, pH o consistencia generan una serie de tiempo que describe su comportamiento, pero estas series no son independientes entre sí.

Cuando ocurre una anomalía crítica, como un corte de hoja, resulta clave analizar cómo se comportan las series de tiempo de cada sensor en instantes previos a este evento, como influyen al resto del sistema y qué sensores se relacionan con más fuerza al corte de hoja, siendo estos últimos los candidatos a ser causas precursoras de la anomalía.

2.2.2. Diebold–Yilmaz

El enfoque propuesto por [Diebold and Yilmaz \(2012\)](#) permite **cuantificar el grado de conectividad bidireccional** entre las variables de un sistema multivariado, posibilitando la identificación de transmisores y receptores de shocks.

Posteriormente, [Demirer et al. \(2018\)](#) y [Barigozzi and Brownlees \(2019\)](#) incorporan el método LASSO (Operador de Reducción y Selección de Mínimos Absolutos) para reducir el espacio de parámetros del modelo VAR y controlar el crecimiento del número de parámetros asociado a sistemas de alta dimensión.

Con el fin de manejar la elevada dimensionalidad de los datos, en este trabajo se utiliza el enfoque propuesto por [Demirer et al. \(2018\)](#) para calcular las medidas de conectividad de Diebold y Yilmaz entre sensores pertenecientes a distintas etapas de la máquina de secado.

Las medidas de conectividad se construyen a partir de la **descomposición de la varianza del error de predicción** del modelo LASSO-VAR. En este contexto, el error de predicción corresponde a la diferencia entre el valor observado de una variable y el valor estimado por el modelo al realizar una predicción a H pasos adelante. La varianza de este error mide la incertidumbre asociada a la predicción futura y permite cuantificar en qué medida dicha incertidumbre es explicada por perturbaciones propias de la variable o por shocks provenientes de otras variables del sistema.

Estas medidas se basan en el marco de modelado VAR generalizado propuesto por [Koop et al. \(1996\)](#) y [Pesaran and Shin \(1998\)](#), denominado KPPS, el cual genera descomposiciones de la varianza invariantes al orden de las variables del sistema. Para $H = 1, 2, \dots$, la descomposición de la varianza del error de predicción a H pasos adelante está dada por:

$$\theta_{ij}^g(H) = \frac{\sigma_{jj}^{-1} \sum_{h=0}^{H-1} (e_i' A_h \Sigma e_j)^2}{\sum_{h=0}^{H-1} (e_i' A_h \Sigma A_h' e_i)}. \quad (2.2.1)$$

donde:

- Σ es la matriz de varianzas y covarianzas del vector de errores ϵ .
- σ_{jj} corresponde a la varianza del error asociado a la j -ésima ecuación.
- e_i es un vector de selección con valor uno en el elemento i -ésimo y ceros en el resto.
- A_h representa la matriz de coeficientes de la representación de media móvil del modelo VAR, que describe la propagación temporal de los shocks.

Dado que al suma de los elementos de cada fila no es igual a 1, $\sum_{j=1}^N \theta_{ij}^g(H) \neq 1$, para evitar esto, normalizamos cada entrada de la matriz de descomposición de varianza:

$$\tilde{\theta}_{ij}^g(H) = \frac{\theta_{ij}^g(H)}{\sum_{j=1}^N \theta_{ij}^g(H)} \quad (2.2.2)$$

Ahora por construcción $\sum_{j=1}^N \tilde{\theta}_{ij}^g(H) = 1$ y $\sum_{i,j=1}^N \tilde{\theta}_{ij}^g(H) = N$

2.2.2.1. Matriz de Conectividad

Con las medidas de spillover normalizadas en 2.2.2, se define a continuación la matriz de conectividad direccional en 2.2.1.

Definimos los x_N con $n \in \{1, \dots, N\}$ como todas las variables del sistema, de las cuales se busca cuantificar la conectividad entre ellas.

El bloque superior izquierdo principal $N \times N$ contiene las descomposiciones de la varianza entre cada par de x_n . Llamaremos a este bloque **matriz de descomposición de varianza** y la denotamos por $S^H = [S_{ij}^H]$.

	x_1	x_2	\dots	x_N	FROM
x_1	S_{11}^H	S_{12}^H	\dots	S_{1n}^H	$\sum_{j=1}^N S_{1j}^H, j \neq 1$
x_2	S_{21}^H	S_{22}^H	\dots	S_{2N}^H	$\sum_{j=1}^N S_{2j}^H, j \neq 2$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
x_N	S_{N1}^H	S_{N2}^H	\dots	S_{NN}^H	$\sum_{j=1}^N S_{Nj}^H, j \neq N$
TO	$\sum_{i=1}^N S_{i1}^H, i \neq 1$	$\sum_{i=1}^N S_{i2}^H, i \neq 2$	\dots	$\sum_{i=1}^N S_{iN}^H, i \neq N$	$\frac{1}{N} \sum_{i,j=1}^N S_{ij}^H, i \neq j$

Cuadro 2.2.1: Esquema de matriz de conectividad

La medida de conectividad o influencia de la variable x_j a la variable x_i , es definida como el **Spillover direccional por pares** y se calcula de la forma:

$$S_{ij}^H = S_{i \leftarrow j}^H = \tilde{\theta}_{ij}^g \quad (2.2.3)$$

Definimos el valor de la fila TO como el **Spillover direccional total recibido por la variable x_i de las demás variables x_j** y se calcula de la forma:

$$S_{i \leftarrow \cdot}^H = \frac{\sum_{j=1, j \neq i}^N S_{ij}^H}{N} \quad (2.2.4)$$

El valor de la columna FROM es definido como el **Spillover direccional total de la variable x_j a todas las demás variables x_i** y se calcula como:

$$S_{\cdot \leftarrow j}^H = \frac{\sum_{i=1, i \neq j}^N S_{ij}^H}{N} \quad (2.2.5)$$

Dado que la magnitud de los shocks recibidos y dados pueden ser distintos, se puede saber si una variable es un receptor o transmisor neto de shocks. Definimos el **Spillover direccional neto de la variable** x_i como:

$$S_i^H = S_{\leftarrow i}^H - S_{i\leftarrow}^H. \quad (2.2.6)$$

El **Spillover Total** es el índice de conectividad total del sistema, representa que tan fuerte es la conexión entre todas las variables dependiendo del instante de tiempo, se define:

$$S^H = \frac{1}{N} \sum_{i,j=1, i \neq j}^N S_{ij}^H \quad (2.2.7)$$

2.3. Modelos de Regresión para Detección de Anomalías

En esta sección se describen los modelos de regresión utilizados para predecir el comportamiento de los sensores y detectar anomalías en el proceso de secado. Se incluyen modelos lineales y basados en árboles, así como algoritmos de aprendizaje basado en vecinos y boosting. Se detallan además los hiperparámetros más relevantes de cada modelo.

En el contexto de los modelos de regresión, un **hiperparámetro** corresponde a una variable de configuración del modelo cuyo valor se fija antes del proceso de entrenamiento y no se estima directamente a partir de los datos. Los hiperparámetros controlan la complejidad del modelo, su capacidad de generalización y el compromiso entre sesgo y varianza, influyendo de manera significativa en el desempeño predictivo. En este trabajo, los hiperparámetros de cada modelo se seleccionan mediante validación cruzada temporal, con el objetivo de optimizar la detección temprana de anomalías.

2.3.1. Correlación de Pearson

La correlación de Pearson es una medida estadística que cuantifica la intensidad y dirección de la relación lineal entre dos variables aleatorias continuas. Sea $X = (x_1, x_2, \dots, x_n)$ y $Y = (y_1, y_2, \dots, y_n)$ dos variables con medias \bar{x} y \bar{y} , respectivamente. El coeficiente de correlación de Pearson, denotado por r_{XY} , se define como

$$r_{XY} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (2.3.1)$$

El valor de r_{XY} se encuentra en el intervalo $[-1, 1]$. Valores cercanos a 1 indican una relación lineal positiva fuerte, mientras que valores cercanos a -1 reflejan una relación lineal negativa fuerte. Un valor cercano a 0 sugiere ausencia de relación lineal.

En este trabajo, la correlación de Pearson se calculó como un paso preliminar para cuantificar el grado de dependencia lineal entre los sensores. Este análisis permite identificar variables que presentan relaciones significativas entre sí y con la variable objetivo, facilitando la selección de predictores relevantes para los modelos de regresión utilizados en la detección temprana de anomalías (Pearson (1895)). La estimación de las correlaciones se realizó mediante la función `corr()` de la biblioteca `pandas` en Python.

2.3.2. Regresión Lineal (LR)

La regresión lineal es un modelo estadístico utilizado para estimar la relación entre una variable dependiente continua y un conjunto de variables independientes. La predicción se obtiene mediante:

$$\hat{y} = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \quad (2.3.2)$$

donde β_0, \dots, β_p son los parámetros ajustados mediante mínimos cuadrados ordinarios (Montgomery et al. (2021)).

2.3.3. Support Vector Regression (SVR)

El SVR busca encontrar una función que aproxime la relación entre un conjunto de variables de entrada y una variable de salida continua, minimizando un margen de tolerancia ϵ . Formalmente, dado un conjunto de entrenamiento $D = \{(x^{(i)}, y^{(i)})\}$, el modelo estima:

$$f(x) = w^T x + b \quad (2.3.3)$$

donde w define la dirección del hiperplano y b ajusta su posición [Smola and Schölkopf \(2004\)](#).

2.3.3.1. Hiperparámetros clave

- **C**: penalización por errores fuera del margen ϵ .
- ϵ : ancho del margen de tolerancia donde no se penalizan errores.
- **Kernel**: tipo de función de transformación (lineal, RBF, polinomial).
- **Gamma**: parámetro del kernel RBF que controla la influencia de cada punto de entrenamiento.

2.3.4. Random Forest (RF)

Random Forest es un modelo de ensamble basado en árboles de decisión, introducido por [Breiman \(2001\)](#), que construye múltiples árboles usando subconjuntos aleatorios de predictores y promedia sus predicciones para regresión:

$$\hat{y} = \frac{1}{K} \sum_{k=1}^K h(x, \theta_k) \quad (2.3.4)$$

donde K es el número de árboles y $h(x, \theta_k)$ es la predicción del árbol k ([Liaw and Wiener \(2002\)](#)).

2.3.4.1. Hiperparámetros clave

- **n_estimators**: cantidad de árboles en el bosque.
- **max_depth**: Máxima profundidad, controla la complejidad de los árboles.

2.3.5. K-Nearest Neighbors (KNN)

KNN es un algoritmo no paramétrico basado en instancias. Para regresión, la predicción de un nuevo punto x_0 se calcula como el promedio de los k vecinos más cercanos:

$$\hat{y}_0 = \frac{1}{k} \sum_{i=1}^k y_{(i)} \quad (2.3.5)$$

donde $y_{(i)}$ son los valores de los vecinos más cercanos ([Altman \(1992\)](#)).

2.3.5.1. Hiperparámetros clave

- **k**: número de vecinos a considerar.

2.3.6. XGBoost

XGBoost es un algoritmo de ensamble basado en árboles y gradient boosting, optimizado para velocidad y regularización. Cada árbol intenta corregir los errores de los anteriores, y la predicción final se calcula como:

$$\hat{y}^{(t)} = \sum_{i=1}^T f_i(x), \quad \Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (2.3.6)$$

donde $\Omega(f)$ es el término de regularización, T el número de hojas y w los pesos de las hojas ([Chen and Guestrin \(2016\)](#)).

2.3.6.1. Hiperparámetros clave

- **n_estimators**: Número de árboles.
- **learning_rate**: Tasa de aprendizaje.
- **max_depth**: Máxima profundidad de los árboles.

2.4. Validación de modelos

2.4.1. Validación cruzada

El entrenamiento de los modelos con datos normales se realizó utilizando una técnica de validación cruzada especializada para series temporales, denominada **TimeSeriesSplit**, implementada en la biblioteca **scikit-learn** (Pedregosa et al. (2011)). Esta metodología está diseñada específicamente para conjuntos de datos ordenados cronológicamente, como las series temporales provenientes de los sensores de la máquina secadora.

A diferencia de métodos de validación tradicionales como **KFold**, donde las divisiones se generan de forma aleatoria (rompiendo la estructura temporal de los datos), **TimeSeriesSplit** conserva el orden cronológico en todas sus particiones (**scikit-learn developers** (2024)). En cada iteración, se emplean los primeros k bloques de datos como conjunto de entrenamiento y el bloque $(k + 1)$ como conjunto de validación, de modo que el modelo aprende sobre información pasada y se evalúa sobre el futuro inmediato.

Este enfoque resulta especialmente adecuado en contextos de predicción y monitoreo temporal, donde utilizar datos futuros para el entrenamiento comprometería la validez del modelo y su capacidad de generalización. En este trabajo, se aplicó la técnica con cinco divisiones consecutivas (**n_splits** = 5), lo que permite evaluar de manera progresiva la estabilidad y consistencia del modelo a lo largo del tiempo.

Para asegurar la comparabilidad entre pliegues, se exige que las observaciones estén equiespaciadas en el tiempo. Bajo esta condición, cada conjunto de validación representa una ventana temporal uniforme, permitiendo obtener métricas de desempeño coherentes y comparables entre las distintas divisiones.

2.4.1.1. Cálculo del error

El desempeño de cada modelo de regresión se evaluó mediante el cálculo del error cuadrático medio (MSE) en cada fold de la validación cruzada temporal (`TimeSeriesSplit`). Esta métrica permite cuantificar la capacidad predictiva del modelo sobre datos no vistos y es ampliamente utilizada en contextos de regresión (Montgomery et al. (2021)).

El error se define como:

$$MSE_j = \frac{1}{n_j} \sum_{i=1}^{n_j} (y_{i,j} - \hat{y}_{i,j})^2, \quad (2.4.1)$$

donde,

- j : índice del fold de validación cruzada.
- n_j : número de muestras del conjunto de validación en el fold j .
- $y_{i,j}$: valor real de la muestra i en el fold j .
- $\hat{y}_{i,j}$: valor predicho para la muestra i del fold j .

2.4.2. Validación Hold-out

La validación *hold-out* es una técnica de evaluación que consiste en dividir el conjunto de datos en dos particiones: una destinada al entrenamiento del modelo y otra utilizada exclusivamente para su validación. Este procedimiento permite estimar el desempeño predictivo del modelo sobre datos no vistos, evitando así un ajuste excesivo a la muestra de entrenamiento (James et al. (2013)).

En este trabajo, se empleó la validación *hold-out* para evaluar el rendimiento de los modelos entrenados únicamente con datos considerados normales, aplicándolos posteriormente sobre intervalos reales que contienen eventos anómalos. Esto permite medir la capacidad del modelo para generalizar desde condiciones operacionales estables hacia escenarios donde pueden aparecer desviaciones o fallas.

2.4.2.1. Cálculo del error

El desempeño predictivo de cada modelo se evalúa mediante el Error Cuadrático Medio (MSE), una de las métricas más utilizadas en problemas de regresión debido a su sensibilidad a errores grandes (Montgomery et al. (2021)). Para cada sensor, el MSE se define como:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (2.4.2)$$

donde,

- n : cantidad de observaciones del conjunto de prueba,
- y_i : valor real de la muestra i ,
- \hat{y}_i : predicción del modelo para la muestra i .

El MSE penaliza de forma cuadrática las discrepancias entre valores reales y predichos, lo que lo convierte en una métrica especialmente sensible ante desviaciones abruptas o cambios estructurales en el comportamiento de los datos. En contextos industriales, esta propiedad es particularmente útil para la detección temprana de anomalías, ya que incluso desviaciones pequeñas pero persistentes pueden generar aumentos significativos en el error, señalando posibles fallas en sensores o procesos operacionales.

Capítulo 3

Metodología

3.1. Datos

3.1.1. Fuente de Datos

CMPC Celulosa, una filial del Grupo CMPC, desempeña un papel clave en esta industria tanto a nivel nacional como internacional. Además de su alta capacidad productiva, la empresa se distingue por su compromiso con la sostenibilidad, la innovación y la adopción de tecnologías avanzadas orientadas a la mejora continua de sus procesos. Gracias a estas iniciativas, ha sido reconocida como una de las compañías forestales más sostenibles del mundo, [CMPC \(2024\)](#).

Los datos utilizados en este estudio provienen de la planta CMPC Celulosa Santa Fe, ubicada en la comuna de Nacimiento, Región del Bio-Bio. En particular, la información corresponde a registros operacionales obtenidos desde la máquina secadora FLAKT, unidad donde se lleva a cabo la etapa final del proceso de producción de celulosa.

El conjunto de datos proviene del sistema de control de la planta, el cual recopila mediciones operacionales en tiempo real de múltiples sensores distribuidos estratégicamente a lo largo de la máquina secadora. Estos sensores capturan variables clave del proceso, tales como flujo, presión, temperatura, consistencia, pH, torque, tensión y otras magnitudes relevantes para la operación y control del sistema.

Los sensores se encuentran distribuidos a lo largo de la máquina secadora como muestra la figura 3.1.1, siguiendo la secuencia operativa del proceso de secado, lo

que permite dividir la operación en cinco etapas principales que representan el recorrido de la celulosa desde su entrada hasta la salida final. Cada una de estas etapas agrupa un conjunto de sensores asociados a funciones específicas del sistema, lo que permite caracterizar el comportamiento y las condiciones operativas de cada zona del proceso.

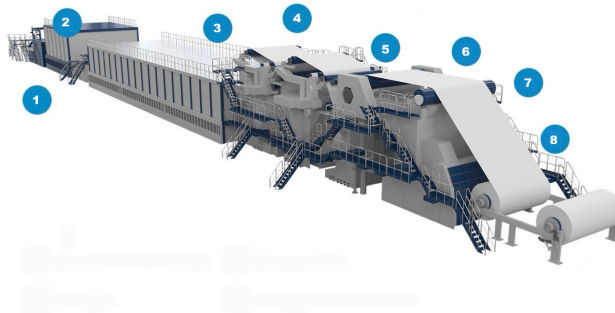


Figura 3.1.1: Máquina secadora FLAKT y distribución de las etapas del proceso.

3.1.2. Sensores

Las mediciones analizadas en este estudio corresponden al período comprendido entre enero y mayo de 2025, durante el cual se registraron de forma continua las variables operacionales provenientes de 96 sensores distribuidos a lo largo de la máquina secadora FLAKT. Cada sensor genera una serie de tiempo con una frecuencia de muestreo de cinco minutos, lo que permite representar con detalle la evolución temporal del proceso, evaluar su dinámica operativa y detectar posibles interacciones o dependencias entre las distintas zonas del sistema.

Cada uno de los 96 sensores cumple una función específica dentro del proceso, registrando variables clave como presiones, caudales, temperaturas, estados de vacío y condiciones mecánicas. La descripción funcional de cada sensor, junto con su ubicación y etapa del proceso, se presenta en las Tablas 3.1.1 y 3.1.2. Esta organización permite analizar de manera segmentada el comportamiento del sistema y estudiar cómo las variaciones locales pueden influir en la dinámica global del proceso de secado.

La máquina secadora se compone de ocho etapas que operan de manera secuencial, donde cada sensor se asocia a una fase específica que registran

variables representativas de su función del proceso productivo. Esta distribución estructurada permite estudiar el comportamiento del sistema de manera segmentada, las interacciones entre las distintas zonas y comprender cómo las variaciones locales pueden influir en el comportamiento global del sistema.

A continuación, se describen las etapas:

1. **Tac Blanca:** Corresponde a la etapa inicial del proceso, donde la pasta es preparada y homogenizada antes de ser enviada hacia el sistema de formación. Los sensores de esta zona miden variables relacionadas con el flujo y la consistencia de la pasta en distintos tanques, además del control de condensados. Esta etapa resulta esencial para asegurar una alimentación estable y uniforme hacia las fases posteriores del proceso.
2. **Depuración:** En esta fase se eliminan impurezas o partículas no deseadas mediante sistemas de harneros primarios, secundarios y terciarios. Los sensores registran diferenciales de presión, permitiendo evaluar la eficiencia del filtrado y el estado operativo de los equipos. Un control adecuado en esta zona garantiza una pasta limpia y homogénea antes de su ingreso a la máquina formadora.
3. **Aproximación:** Actúa como una zona de transición donde se regulan el flujo, nivel, presión y temperatura de la pulpa que ingresa a la caja formadora. Los sensores instalados en esta etapa permiten estabilizar las condiciones de operación, asegurando que la mezcla llegue con las características adecuadas para la formación de la hoja.
4. **Formación:** Esta etapa es una de las más críticas del proceso, ya que aquí la pulpa líquida se transforma en una hoja continua. Los sensores monitorean variables clave como presión, torque, temperatura y pH, que determinan la calidad y uniformidad de la hoja. El control preciso de estos parámetros es fundamental para mantener la estabilidad operativa y evitar defectos durante la formación.
5. **Sistema de Vacío:** Etapa del proceso en la que se realiza la extracción inicial de agua de la hoja mediante la aplicación de vacío. Su función principal es reducir el contenido de humedad antes de la etapa de prensado, mejorando la formación y estabilidad de la hoja. Los sensores asociados a esta etapa miden variables de presión de vacío, niveles y caudales, permitiendo monitorear la

eficiencia del sistema y detectar posibles fallas operacionales.

6. **Prensas:** Etapa en la que se efectúa la extracción mecánica de agua mediante la aplicación de presión entre rodillos. Esta fase contribuye a la compactación de la hoja y al aumento de su resistencia estructural, preparándola para el secado térmico posterior. Los sensores de esta zona registran variables como presiones de prensado, torques de los rodillos, temperaturas y condiciones de lubricación, las cuales son críticas para la operación segura y eficiente del proceso.
7. **Secador:** En la etapa de secado, la hoja pasa a través de secciones con aire y vapor caliente que reducen su contenido de humedad hasta los niveles requeridos. Los sensores miden temperatura, presión, flujo de vapor, apertura de válvulas y tensiones en la hoja. Estas variables permiten optimizar la eficiencia energética del proceso y mantener la calidad del producto final.
8. **Torre de Recorte:** Corresponde a una etapa del proceso destinada a la recuperación de los recortes generados durante el conformado y dimensionamiento de la hoja de celulosa. Estos recortes son recolectados y conducidos a la torre de recorte, donde son acondicionados para su posterior reincorporación al proceso de secado. Los sensores asociados a esta etapa miden variables relacionadas con el nivel y la consistencia del material recuperado, así como las corrientes eléctricas de las bombas utilizadas en la recirculación de la pasta, contribuyendo a la eficiencia y estabilidad global del sistema.

Cuadro 3.1.1: Descripción sensores

Etapa	ID	Sensor	Descripción
TAC BLANCA	1	25ONIC008.MEAS	CNTL CONS SALIDA TK1 PBL
	2	25ONV008.OUT	CNTL CONS SALIDA TK1 PBL
	3	25ONIC028.MEAS	CNTL CONS SALIDA TK2 PBL
	4	25ONV028.OUT	CNTL CONS SALIDA TK2 PBL
	5	25ONIC908.MEAS	SALIDA TK CONTROL CONDENSADO
	6	25OFIC012.MEAS	CNTL DE FLUJO TK2 P BL A TK MIX
	7	25OFIC032.MEAS	CNTL DE FLUJO TK2 P BL A TK MIX
	8	25OFIC912.MEAS	TK3 A TK MIX CNTL DE FLUJO
	9	25ONT283.MEAS	CONSISTENCIA SALIDA TORRE RECORTE
	10	25ONV283.OUT	SALIDA TK RECORTE CONTROL CONSIS
	11	25OFT042.MEAS	FLUJO PASTA DE TORRE DE RECORTE
	12	25OFV042.OUT	TK RECORTE A CONTROL FLUJO
TORRE RECORTE	13	25OLT281.MEAS	NIVEL TK RECORTE
	14	25ONIC046.MEAS	CONSISTENCIA TK MEZCLA
	15	2502300101II.PNT	CORRIENTE BOMBA TK MEZCLA 2300101
DEPURACIÓN	16	250PDIT066.PNT	DIFERENCIAL DE PRESION HARNERO PRIMARIO 1
	17	250PDIT086.PNT	DIFERENCIAL DE PRESION HARNERO PRIMARIO 2
	18	250PDIT106.PNT	DIFERENCIAL DE PRESION HARNERO PRIMARIO 3
	19	250PDIT126.PNT	DIFERENCIAL DE PRESION HARNERO PRIMARIO 4
	20	250PDIT146.PNT	DIFERENCIAL DE PRESION HARNERO SECUNDARIO
	21	250PDIT166.PNT	DIFERENCIAL DE PRESION HARNERO TERCARIO
APROXIMACIÓN	22	25ONIC293.MEAS	CONSISTENCIA TK MAQUINA
	23	2502304001II.PNT	CORRIENTE BOMBA 2304001
	24	250PIT332.PNT	AGUA CONS Y WAC TK PRESION WAC
	25	25OFIC300.MEAS	FLUJO PASTA TK MAQUINA
	26	25OLT291.MEAS	NIVEL TK MAQUINA
	27	25OFIC303.MEAS	A TELA FORMADORA
	28	25OFIC302.MEAS	FLUJO TRA CAJON CA
	29	2502304501SI.PNT	VELOCIDAD BOMBA FAN
	30	2502304501II.PNT	CORRIENTE BOMBA FAN
	31	250TIC323.MEAS	FLUJO RIN CAJON CA CONTROL TEMP
FORMACIÓN	32	25OPT336.MEAS	PRESION PASTA BOMBA FAN
	33	250COMP332.RIO2	pH CAJA DE ENTRADA
	34	250TT323.MEAS	TEMPERATURA PASTA CAJA ENTRADA
	35	250JET WIRE.R004	JET WIRE
	36	250PIT333.PNT	PRESION CAJON C LADO SERVC
	37	250PIT334.PNT	PRESION CAJON C LADO ACC
	38	250PIT354.PNT	CORTE DE BORDES PRESION FH
	39	2112401IT.PNT	TORQUE ACCIONAMIENTO TELA SUPERIOR
	40	2112501IT.PNT	TORQUE ACCIONAMIENTO RODILLO COUCH
	41	2112601IT.PNT	TORQUE ACCIONAMIENTO TELA INFERIOR
	42	250PIC602.MEAS	PRESION CAJA DE VAPOR
	43	250FIT601.PNT	FLUJO VAPOR DE CAJA DE VAPOR
	44	250TIC604.MEAS	FLUJO VAPOR DE CAJA DE VAPOR
	45	250TIT609.PNT	TEMPERATURA VAPOR CAJA DE VAPOR

Cuadro 3.1.2: Descripción sensores

Etapa	ID	Sensor	Descripción
SISTEMA DE VACIO	46	250PIT668.PNT	INFERIOR 1RA PRESSMEDIDA PRESION
	47	250PIT669.PNT	2DA INFERIOR MEDIDA PRESION
	48	250PIT640.PNT	SEPARACION MEDIDA PRESION
	49	250PIT641.PNT	VACIO CAJA DOBLE MEDIDA PRESION
	50	250PIT642.PNT	VACIO CAJA DOBLE MEDIDA PRESION
	51	250PIT643.PNT	VACIO CAJA TRIPLE MEDIDA PRESION
	52	250PIT671.PNT	VACIO CAJA TRIPLE MEDIDA PRESION
	53	250PIT670.PNT	VACIO CAJA TRIPLE MEDIDA PRESION
	54	250PIT663.PNT	RODILLO SUCC TELA MEDIDA PRES
	55	250PIC726.MEAS	ZONA DE RETENCION CONTROL PRES
	56	250PIT693.PNT	ZONA PRESION MEDIDA PRES
	57	250PIT723.PNT	CAJA PAÑO SUPERIOR 1RA MEDIDA PRES
	58	250PIT724.PNT	CAJA PAÑO SUPERIOR 1RA MEDIDA PRES
	59	250PIT634.PNT	CAJA PAÑO INF 1RA MEDIDA PRES
60	250PIT753.PNT	CAJA PAÑO SUPERIOR 2DA MEDIDA PRES	
61	250PIT754.PNT	CAJA PAÑO INF 2DA MEDIDA PRES	
PRENSAS	62	2114001IT.PNT	TORQUE ACCIONAMIENTO RODILLO PICK-UP
	63	2115601IT.PNT	TORQUE ACCIONAMIENTO 2a PRENSA
	64	2116901IT.PNT	TORQUE ACCIONAMIENTO 1 - PRENSA SHOE PRESS
	65	2116902IT.PNT	TORQUE ACCIONAMIENTO 2 - PRENSA SHOE PRESS
	66	250PIT364A.PNT	LUMP BREAKER CARGANDO ZONA
	67	250PIT364B.PNT	LUMP BREAKER CARGANDO ZONA
	68	250PIT365A.PNT	NIP 1RA PRES CARGANDO ZONA
	69	250PIT365B.PNT	NIP 1RA PRES CARGANDO ZONA
	70	250PIT366A.PNT	NIP 2DA PRES CARGANDO ZONA
	71	250PIT366B.PNT	NIP 2DA PRES CARGANDO ZONA
	72	250PT726.MEAS	PRESION VACIO ZONA PICK-UP
	73	250TIT497.PNT	SIST LUBRICACION CNTL PRESION
	74	250TIT478.PNT	TEMP. ACEITE ENFR. Y LUBRIC.
	SECADOR	75	250 TEN IND.PNT
76		250PIC514.MEAS	PRESION DE VAPOR AL SECADOR
77		250FIC511.MEAS	FLUJO DE VAPOR AL SECADOR
78		250TIT512.PNT	TEMPERATURA VAPOR AL SECADOR
79		250PV514A.OUT	CONTROL ABERTURA VALVULA VAPOR 514A
80		250PV514B.OUT	CONTROL ABERTURA VALVULA VAPOR 514B
81		250PV514C.OUT	CONTROL ABERTURA VALVULA VAPOR 514C
82		250TIC530.MEAS	TEMPERATURA SECADOR
83		250PIT550A.PNT	PRESION VAPOR SECADOR
84		250PIT550.MEAS	PRESION INTERNA AIRE SECADOR
85		250HV523.OUT	DAMPER 1 - INYECCION AIRE SECADOR
86		250HV533.OUT	DAMPER 2 - INYECCION AIRE SECADOR
87		250HV543.OUT	DAMPER 3 - INYECCION AIRE SECADOR
88		250HV527.OUT	DAMPER 1 - EXTRACCION AIRE SECADOR
89		250HV537.OUT	DAMPER 2 - EXTRACCION AIRE SECADOR
90		250HV547.OUT	DAMPER 3 - EXTRACCION AIRE SECADOR
91		250HV557.OUT	DAMPER 4 - EXTRACCION AIRE SECADOR
92		250LIC518.MEAS	NIVE SEPARADOR DE CONDENSADO
93	250TIT516B.PNT	TEMPERATURA VAPOR FLASH TK CONDENSADO	
94	250BASEACT.MEAS	GRAMAJE	
95	250MOIST.MEAS	HUMEDAD DCI BLOCK TYPES MOIST	
96	2502128201.R002	VELOCIDAD PRENSA EXTRACCION SECADOR	

3.1.3. Tiempos perdidos

El registro de tiempos perdidos corresponde a un archivo operativo mantenido por los operarios de la máquina secadora de celulosa, en el cual se documentan manualmente los intervalos de tiempo en que el proceso detiene su funcionamiento normal o presenta condiciones anómalas. Este registro constituye una fuente de información complementaria al sistema automático de medición, ya que permite identificar de forma precisa los periodos de inactividad o alteración del proceso y relacionarlos con las posibles causas que los originan.

En dicho documento se registran los eventos que interrumpen o afectan la continuidad operacional de la máquina, junto con su fecha, duración y motivo asociado. Estas interrupciones pueden deberse tanto a fallas de proceso o equipos, como a actividades de mantenimiento o ajustes programados, y resultan fundamentales para contextualizar los comportamientos observados en las series temporales de los sensores.

Los eventos registrados incluyen:

- **Corte de hoja:** detención del proceso por la ruptura o desprendimiento de la hoja de celulosa.
- **PNP (Parada de Planta No Programada Forzada):** interrupciones imprevistas que obligan a detener la máquina por motivos operativos, mecánicos o de seguridad.
- **PP (Parada de Planta Programada):** detenciones planificadas para realizar labores de mantenimiento, limpieza o ajustes preventivos.
- **Fuera de servicio:** periodos en que el sistema o sus componentes no están operativos, ya sea por reparaciones o condiciones externas al proceso.
- **PGP (Parada General Programada):** detenciones coordinadas a nivel de planta que afectan múltiples unidades de producción simultáneamente.
- **Sensor de tensión:** Evento asociado a lecturas de tensión inferiores a 50, las cuales indican una detención de la máquina, ya sea producto de un corte de hoja o de una parada programada. Estos eventos reflejan una pérdida abrupta de tensión en la hoja y constituyen un indicador operativo relevante para la identificación de interrupciones en el proceso.

- **Sensor de velocidad menor a 130:** evento que refleja una desaceleración significativa del sistema, usualmente vinculada a una condición de operación inestable o a un procedimiento de parada controlada.

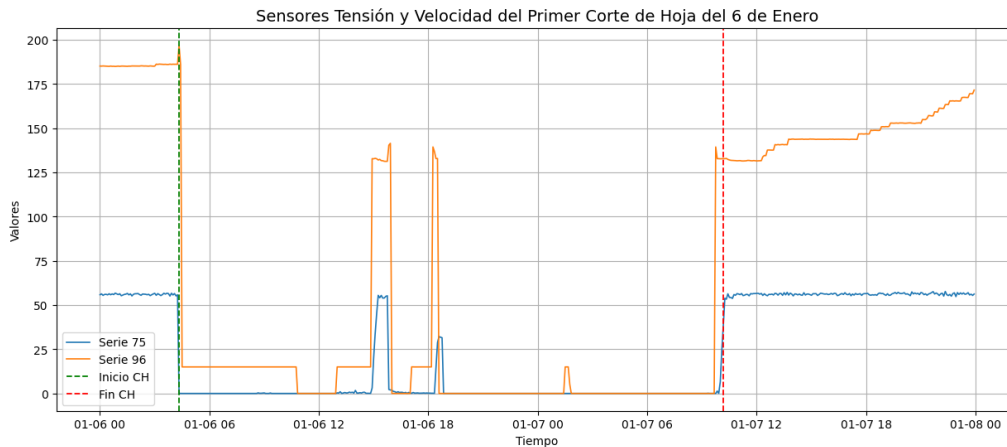


Figura 3.1.2: Sensor 75 de Tensión y 96 de Velocidad cercano al Corte de Hoja 1.

El análisis de estos registros permite delimitar los periodos de interés para el estudio, identificar los instantes previos a las fallas y contrarrestar sus efectos con la implementación de acciones preventivas y correctivas orientadas a reducir el impacto operativo y evitar la recurrencia de fallas similares.

3.1.4. Bitácora Cortes de hoja

A partir del registro de tiempos perdidos mantenido por los operarios de la máquina secadora, se extrajeron específicamente los eventos clasificados como cortes de hoja (CH) para su análisis detallado. Estos eventos representan las interrupciones más críticas dentro del proceso de secado, ya que implican la detención completa de la línea y generan impactos tanto operativos como productivos.

Cada evento de corte de hoja se encuentra documentado con información que incluye la fecha y hora de inicio y término, el motivo reportado y la etapa del proceso afectada. Este nivel de detalle permite identificar las condiciones operativas previas al evento, analizar las posibles causas que lo originaron y evaluar la recurrencia de patrones o variables precursoras detectadas en los sensores del sistema.

El análisis de esta bitácora tiene como propósito comprender la dinámica de los cortes desde una perspectiva temporal y causal, observando cómo evolucionan las variables de proceso antes, durante y después de cada evento. De esta forma, se busca establecer relaciones entre el comportamiento de los sensores y la ocurrencia de los cortes de hoja, identificando posibles indicadores tempranos de falla que permitan anticipar o mitigar futuros eventos.

A partir del Cuadro 3.1.3, se definen y describen los distintos eventos de corte de hoja registrados durante el período de estudio.

Se observa que los cortes CH-5, CH-6 y CH-7 ocurrieron con poca separación temporal entre sí, lo que sugiere la existencia de condiciones operativas similares durante ese periodo. Esta recurrencia resulta particularmente relevante para el estudio, ya que permite evaluar la presencia de patrones comunes en las variables del proceso y comparar su comportamiento previo a cada evento.

Cuadro 3.1.3: Cortes de hoja 2025

n°	Inicio CH	Fin CH	Duración	CH previo	Motivo
1	2025-01-06 04:20:00	2025-01-07 10:10:00	05:50:00	-	Caída de pasta o condensado sobre hoja y falla enclavamiento
2	2025-01-15 23:00:00	2025-01-16 05:45:00	06:45:00	8.54 días	Por aseo en formador
3	2025-02-09 23:00:00	2025-02-11 08:05:00	32:05:00	23.7 días	(fs secado) Evento: Cambio tela inferior por rotura Modo falla: Se desprende lámina de porta labio inferior Causa: Lámina no corresponde al diseño del orta labio inferior, otro proveedor
4	2025-02-11 15:25:00	2025-02-11 19:10:00	02:45:00	0.3 días	salida 2da prensa
5	2025-03-01 17:30:00	2025-03-01 22:30:00	05:00:00	17.93 días	salida 2da prenda lado accionamiento. Contaminación agua mixta y tibia utilizada en chorros formateadores
6	2025-03-02 06:30:00	2025-03-02 13:05:00	06:35:00	0.62 días	Salida 2da prensa lado accionamiento. Contaminación agua mixta y tibia utilizada en chorros formateadores
7	2025-03-03 08:05:00	2025-03-03 13:10:00	05:05:00	0.63 días	Contaminación agua mixta y tibia utilizada en chorros formateadores
8	2025-03-12 10:15:00	2025-03-13 15:30:00	05:15:00	9.93 días	Contaminación agua mixta y tibia utilizada en chorros formateadores
9	2025-03-28 00:50:00	2025-03-28 23:45:00	22:55:00	14.43 días	Bomba Fan modo falla: pérdida aislamiento motor eléctrico causa: degradación aislamiento por alta temperatura. etapa potencial VDF con falla de alimentación a motor
10	2025-04-03 18:50:00	2025-04-04 00:45:00	05:55:00	6.74 días	Bajo nivel TK máquina causa: descontrol de flujos de harneros
11	2025-04-15 08:50:00	2025-04-15 15:10:00	06:20:00	11.34 días	en secador causas: rotula de accionamiento desacoplada, pérdida de vacío en polin pick up
12	2025-04-20 02:20:00	2025-04-20 09:05:00	06:45:00	4.45 días	en formador
13	2025-05-02 22:50:00	2025-05-03 03:45:00	04:55:00	12.55 días	problema agua con ácido. Observable en tag de pH caja de entrada. Problema empieza una hora previa al corte de hoja, detectable
14	2025-05-23 18:35:00	2025-05-24 12:35:00	18:00:00	20.60 días	salida 3ra prensa. Acumulación de pitch en polin nipco

3.1.5. Serie Binaria CH

A partir de la información contenida en la bitácora de cortes de hoja, se construyó una serie de tiempo binaria que representa la ocurrencia de estos eventos a lo largo del período de estudio. En esta serie, el valor 1 indica la presencia de un evento de corte de hoja (CH), mientras que el valor 0 corresponde a condiciones de operación normal del proceso.

La transformación del registro de eventos en una serie binaria permite incorporar la variable de corte de hoja dentro del marco de análisis temporal junto con las demás series de sensores y facilita la identificación automática de ventanas temporales asociadas a cada evento. De esta forma, es posible estudiar su relación dinámica y causal con las variables operacionales del sistema, evaluando qué sensores o grupos de sensores presentan comportamientos que preceden o anticipan la ocurrencia de un corte.

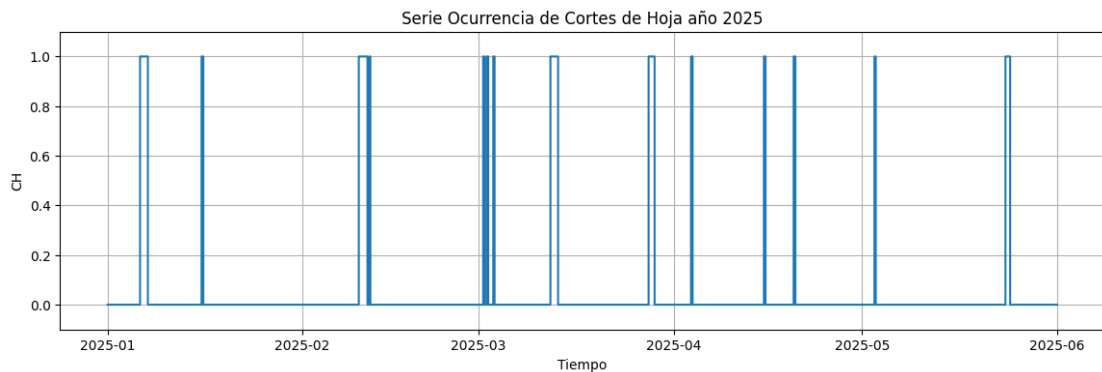


Figura 3.1.3: Serie de tiempo de los cortes de hoja 2025

3.2. Análisis de conectividad

3.2.1. Metodología Spillovers

Para cada evento de corte de hoja se aplicó la metodología de Spillovers propuesta por Diebold y Yilmaz, con el objetivo de analizar la conectividad entre las variables medidas por los sensores del proceso y la ocurrencia del evento de corte de hoja. Esta metodología permite cuantificar las interrelaciones dinámicas entre las series

temporales, identificando qué variables influyen en otras dentro del sistema y, en particular, cuáles presentan una influencia significativa hacia la variable asociada al corte de hoja (CH).

Por cada análisis se consideró la data de 12 horas, que abarcan 6 horas previas y 6 horas posteriores al momento en que ocurre el corte de hoja. Dado que las series presentan una frecuencia de muestreo de cinco minutos, por lo que contiene 144 observaciones por variable, lo que permite capturar tanto las condiciones que anteceden al evento como las fluctuaciones que ocurren durante y después de éste.

Adicionalmente, se incorporó al análisis una serie binaria de ocurrencia de corte de hoja (CH), construida a partir de los registros operacionales. Esta serie adopta el valor 1 desde tres horas antes del inicio del evento hasta su término, y 0 en el resto del tiempo. Su inclusión permite evaluar de forma explícita la conectividad de los sensores hacia el evento del corte de hoja.

3.2.2. Preprocesamiento

3.2.2.1. Tratamiento de series constantes

Para el análisis de conectividad se aplicó un esquema de ventanas móviles de 24 observaciones, equivalentes a 2 horas de operación, desplazadas en intervalos de 5 minutos. Este enfoque permite capturar la evolución temporal de las relaciones entre sensores de forma dinámica.

Durante este proceso, se detectó que algunas series permanecían constantes dentro de ciertas ventanas, lo que puede generar inestabilidad numérica o impedir el ajuste adecuado de los modelos VAR utilizados en la estimación de conectividad. Para evitar este problema, las series identificadas como constantes fueron ajustadas mediante la adición de ruido gaussiano de muy baja magnitud ($\sigma = 1e-10$). Este ruido se aplicó exclusivamente a las series afectadas y en una escala suficientemente pequeña como para no alterar significativamente los valores originales, garantizando que su efecto sobre los resultados fuera despreciable.

De este modo, se aseguró que todas las series presentaran variabilidad mínima dentro de cada ventana temporal, permitiendo su inclusión en el análisis de manera consistente.

3.2.2.2. Estacionariedad

Posteriormente, se verificó la estacionariedad de las series temporales, requisito fundamental para la aplicación del paquete `ConnectednessApproach`. Para ello, se utilizó en R el test de Dickey-Fuller Aumentado (ADF), evaluando la hipótesis nula de presencia de raíz unitaria.

Las series con un p-value menor a 0.05 se consideraron no estacionarias, por lo que fueron transformadas mediante diferenciación de primer orden. Luego, el test ADF se aplicó nuevamente a las series diferenciadas, confirmando que todas alcanzaron un comportamiento estacionario.

Con este procedimiento, se garantizó que todas las series estuvieran en la forma adecuada para su análisis mediante la metodología de conectividad basada en modelos VAR, asegurando la validez estadística de los resultados obtenidos.

3.2.3. Ranking de sensores con mayor conectividad hacia CH

A partir de los resultados obtenidos mediante la metodología de spillovers de Diebold y Yilmaz, se evaluaron las interrelaciones entre todos los sensores y la serie binaria asociada al evento de corte de hoja (CH). Si bien el análisis completo de conectividad permite comprender las dependencias globales del sistema, en este estudio se pone especial énfasis en las conexiones dirigidas hacia la serie CH, con el objetivo de identificar los sensores con mayor influencia potencial en la ocurrencia de un corte de hoja.

Para ello, se considera la ecuación 2.2.3 particularizada para el caso $i=CH$:

$$S_{CH \leftarrow j}^H = \tilde{\theta}_{CH j}^g \quad (3.2.1)$$

A partir de estos valores se construyó un ranking de sensores con mayor conectividad hacia CH, el cual permite identificar aquellos sensores que exhiben una mayor capacidad de transmisión de información o influencia hacia el evento de corte. Este enfoque facilita distinguir los posibles sensores precursores, es decir, aquellos cuyas variaciones anteceden y posiblemente contribuyen a la ocurrencia de un corte de hoja.

3.3. Entrenamiento con datos normales

3.3.1. Construcción de Máscaras Booleanas

El objetivo de esta etapa es entrenar los modelos utilizando únicamente datos correspondientes al comportamiento normal del proceso. De esta forma, al evaluar el modelo con datos reales —que pueden incluir anomalías o condiciones no ideales— se espera que el error aumente significativamente, reflejando la capacidad del modelo para detectar desviaciones respecto a la operación normal.

Para asegurar que el entrenamiento se realizó exclusivamente con datos de operación estable, se implementó una **máscara booleana** que distingue entre periodos normales y periodos afectados por eventos anómalos o condiciones fuera de régimen.

Esta máscara fue construida a partir de un archivo interno de la planta denominado “*Tiempos Perdidos*”, cada vez que se identifica la ocurrencia de alguno de estos eventos, el instante correspondiente se marca como **FALSE** en la máscara, indicando que se trata de un periodo anómalo no apto para el entrenamiento. Además, para evitar que el modelo aprenda patrones asociados a las transiciones entre estados normales y anormales, se excluyen también los datos del día anterior y posterior a cada evento, marcándolos igualmente como **FALSE**.

El resto de los datos, considerados representativos del funcionamiento normal del proceso, se marcan como **TRUE** y constituyen la base del entrenamiento.

Con el fin de visualizar estos periodos, se genera un gráfico binario donde el valor 1 indica condiciones normales (**TRUE**) y el valor 0 indica condiciones anómalas (**FALSE**).

La Figura 3.3.1 muestra la visualización del año 2025 de esta máscara, donde los intervalos en verde corresponden a los periodos clasificados como normales y, por tanto, utilizados para el entrenamiento de los modelos.

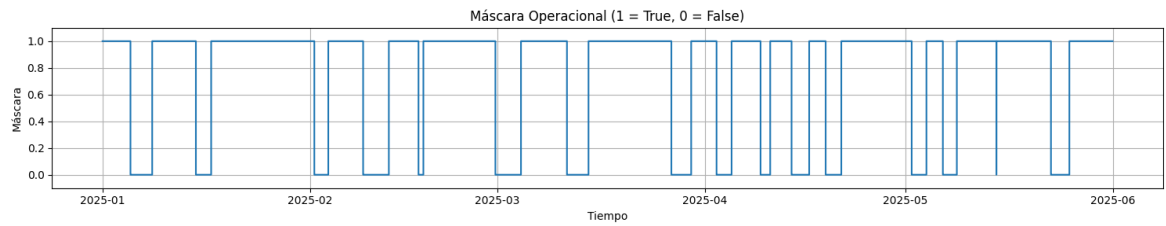


Figura 3.3.1: Máscara operacional para el año 2025

3.3.2. Preprocesamiento

Antes de aplicar los modelos de regresión para la detección de anomalías, fue necesario realizar un preprocesamiento de las series temporales con el objetivo de reducir la redundancia y mejorar la capacidad de interpretación del modelo. Este paso resulta fundamental en entornos industriales, donde los sensores registran múltiples señales fuertemente correlacionadas entre sí, lo que puede conducir a modelos sobreajustados o a interpretaciones erróneas de las dependencias entre variables.

El análisis evidenció que, en la mayoría de los casos, los eventos de corte de hoja se originan en las etapas iniciales del proceso de producción, mientras que las variaciones observadas en etapas posteriores suelen corresponder a efectos derivados o manifestaciones tardías del problema original. La Figura 3.3.2 muestra el flujo general de la celulosa dentro de la máquina de secado, ilustrando el recorrido del material a través de las diferentes etapas del proceso.

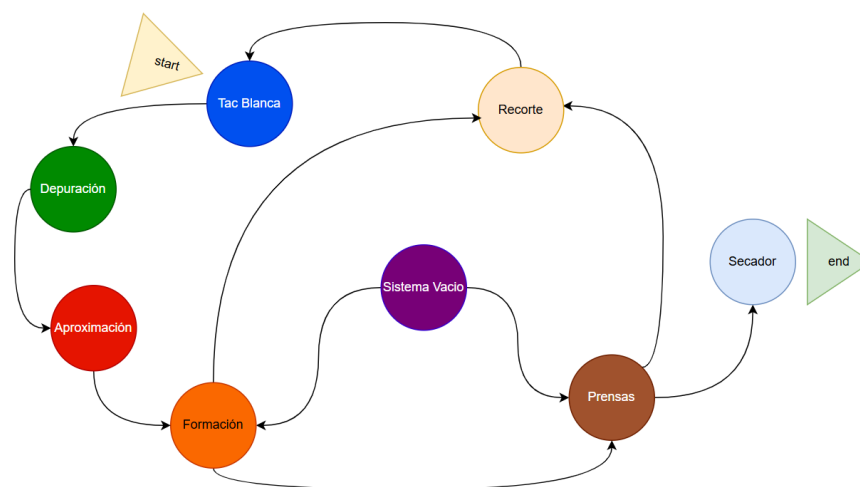


Figura 3.3.2: Mapa de etapas de la máquina secadora

Con base en esta información, se definió un orden de prioridad para la eliminación de variables, privilegiando aquellas ubicadas más próximas al inicio de la máquina, debido a su mayor relevancia causal en la generación de anomalías. En el Cuadro 3.3.1 se presenta dicho orden de prioridad, donde el nivel 1 corresponde a los sensores con mayor relevancia (mantención prioritaria), mientras que el nivel 7 agrupa a los sensores que serán eliminados en caso de alta correlación con otros.

Prioridad	Etapa
1	Tac Blanca
2	Depuración
3	Aproximación
4	Formación
5	Prensas
6	Recorte
7	Secador

Cuadro 3.3.1: Orden de prioridad por cercanía al origen de la máquina

En resumen, cuando dos sensores presentaron una alta correlación, se conservó aquel asociado a una etapa más cercana al origen del proceso. Este criterio no solo permitió reducir la colinealidad entre variables, sino que además aseguró que la selección de características estuviera alineada con el conocimiento del proceso productivo, práctica recomendada en aplicaciones industriales.

3.3.3. Entrenamiento con time series split

Para cada evento de corte de hoja (CH), se analiza la semana inmediatamente anterior al suceso, comprendida entre lunes y domingo. A partir de la máscara booleana descrita en la sección anterior, se identifican los intervalos de dicha semana que corresponden a condiciones operativas normales (**TRUE**).

Las máscaras booleanas de cada semana previa de entrenamiento y su posterior testeó se encuentran en la sección del Anexo A2.

Si toda la semana es clasificada como normal, se utiliza el intervalo completo; en caso contrario, se selecciona únicamente el subintervalo continuo que se mantenga en estado **TRUE**. Este intervalo constituye la base de entrenamiento del modelo correspondiente a ese evento.

A partir de esta información, se diseña un esquema de análisis semanal e individual

por sensor, en el cual cada sensor del sistema es modelado de manera independiente.

El objetivo es que cada modelo aprenda la dinámica temporal propia del sensor bajo condiciones normales, permitiendo posteriormente detectar desviaciones locales o globales asociadas a comportamientos anómalos.

El entrenamiento se realiza utilizando el método `TimeSeriesSplit`, una técnica de validación cruzada específica para series temporales. A diferencia de la validación cruzada tradicional, que divide los datos de forma aleatoria, este método conserva el orden temporal de las observaciones, evitando el uso de información futura para predecir el pasado.

En la práctica, el conjunto de entrenamiento se divide en múltiples particiones crecientes: en cada iteración, el modelo se entrena sobre un bloque inicial de datos y se valida sobre el siguiente segmento temporal. Este procedimiento permite evaluar la estabilidad del modelo y su capacidad de generalización ante nuevas condiciones dentro del mismo régimen operativo.

De este modo, el esquema `TimeSeriesSplit` no solo mejora la robustez del entrenamiento, sino que también proporciona una estimación más realista del desempeño predictivo de los modelos bajo un contexto temporal, respetando la naturaleza secuencial del proceso industrial.

3.3.4. Cálculo MSE

El desempeño de cada modelo se evalúa mediante el Error Cuadrático Medio, calculado para cada sensor, evento de corte de hoja y fold de validación cruzada.

Este error se define como:

$$MSE_{j,m}^{CH,s} = \frac{1}{n_j^{CH}} \sum_{i=1}^{n_j^{CH}} (y_{i,j}^{CH,s} - \hat{y}_{i,j}^{CH,s})^2, \quad (3.3.1)$$

donde,

- j : índice del fold de validación cruzada.
- n_j^{CH} : número de muestras del conjunto de prueba en el fold j .
- $y_{i,j}^{CH,s}$: valor real de la muestra i en el fold j para el sensor s .

- $\hat{y}_{i,j}^{CH,s}$: valor predicho por el modelo m para la muestra i del fold j .
- m : modelo de regresión empleado (KNN, LR, RF, SVR o XGB).

Este indicador cuantifica la desviación promedio entre los valores observados y los predichos por el modelo, siendo especialmente adecuado para evaluar tareas de reconstrucción o predicción continua. Un menor valor de $MSE_{j,m}^{CH,s}$ indica que el modelo logra representar de forma más precisa la dinámica normal del sensor.

3.3.5. Selección de hiperparámetros

El cálculo del error se repite para cada combinación de hiperparámetros definida en la Tabla 3.3.2. La búsqueda se realizó dentro de un rango acotado para cada modelo, con el propósito de equilibrar la capacidad de aprendizaje y la generalización, evitando tanto el sobreajuste como la pérdida de capacidad predictiva.

Cuadro 3.3.2: Posibles parámetros

	1	2	3	4	5	6	7	8
KNN	n_n= 3	n_n= 5	n_n= 7	-	-	-		
LR	-	-	-	-	-	-		
RF	n_e= 50 max_d= None	n_e= 50 max_d= 10	n_e= 100 max_d= None	n_e= 100 max_d= 10	-	-		
SVR	kernel= rbf C= 1 gamma= scale	kernel= rbf C= 1 gamma= auto	kernel= rbf C= 10 gamma= scale	kernel= rbf C= 10 gamma= auto	-	-		
XGB	n_e= 50 max_d= 3 l_r= 0.01	n_es= 50 max_d= 3 l_r= 0.1	n_e= 50 max_d= 5 l_r= 0.01	n_e= 50 max_d= 5 l_r= 0.1	n_e= 100 max_d= 3 l_r= 0.01	n_e= 100 max_d= 3 l_r= 0.1	n_e= 100 max_d= 5 l_r= 0.01	n_e= 100 max_d= 5 l_r= 0.1

En la Figura 3.3.3 se presenta un ejemplo del comportamiento del error MSE para el modelo XGBoost bajo distintas combinaciones de hiperparámetros. En cada sensor se selecciona el conjunto de parámetros que produce el menor error promedio durante el entrenamiento en condiciones normales. Con esta configuración óptima se realiza luego la fase de testeo, utilizando los datos reales asociados a cada evento de corte de hoja.

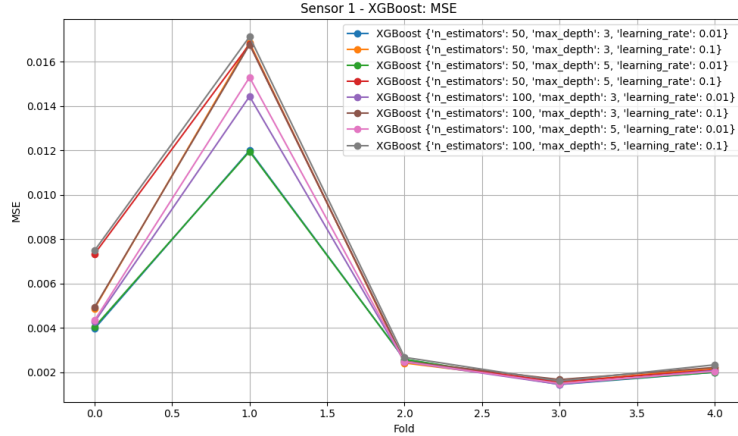


Figura 3.3.3: Error de reconstrucción (MSE) del modelo XGBoost con diferentes combinaciones de parámetros, evaluado por fold

Finalmente, este esquema de validación cruzada temporal constituye la base del proceso de selección de hiperparámetros para cada modelo de regresión individual. En cada sensor se evaluaron múltiples configuraciones —como número de estimadores, profundidad máxima, tasa de aprendizaje o parámetros de regularización— y se seleccionó la que minimizó el MSE promedio entre los folds. De esta manera, se asegura un ajuste equilibrado entre precisión, estabilidad y capacidad de generalización, garantizando que los modelos representen fielmente el comportamiento normal del proceso y puedan detectar desviaciones anómalas de manera robusta.

3.3.6. Definición de umbral de detección

Una vez determinados los hiperparámetros óptimos, se establece un **umbral máximo de error normal** para cada sensor y modelo, con el objetivo de distinguir comportamientos normales de posibles anomalías.

Durante la fase de entrenamiento, se calculan los valores de MSE correspondientes a cada *fold* de la validación cruzada, considerando únicamente los datos bajo condiciones normales. Sea $MSE_{j,m}^{CH,s}$ el error cuadrático medio del fold j , modelo m y sensor s . Para cada sensor s y modelo m , se define el MSE máximo como el valor más alto observado entre todos los folds:

$$Umbral_{MSE_{\max}}^{m,s} = \max_j \left(MSE_{j,m}^{CH,s} \right), \quad (3.3.2)$$

donde $j = 1, \dots, k$ indica los k folds de la validación cruzada. Este valor representa el peor desempeño esperado del modelo bajo condiciones estables y sirve como referencia para la detección de anomalías.

3.4. Entrenamiento con datos reales

3.4.1. Validación hold out

La hipótesis central de este análisis sostiene que un modelo entrenado exclusivamente con datos normales presentará un incremento significativo en el error cuando la dinámica del proceso se vea alterada por una anomalía real. Para formalizar esta detección, se definió un **umbral de normalidad** para cada sensor, correspondiente al mayor valor de MSE observado durante la validación cruzada con datos normales. Si el error observado sobrepasa dicho umbral, se considera que el sensor ha experimentado una desviación respecto de su comportamiento esperado.

Este enfoque permite no solo detectar la ocurrencia de un evento anómalo, sino también identificar los sensores con mayores desviaciones, los cuales pueden actuar como **posibles causas precursoras del corte de hoja**. En consecuencia, el sistema se plantea como un mecanismo de alerta temprana basado en el error de predicción, capaz de anticipar condiciones anómalas en las señales del proceso.

El procedimiento general seguido se describe a continuación:

1. Para una semana determinada (W), se seleccionan los datos normales según la máscara booleana. Con estos datos se entrena un modelo de regresión por sensor, utilizando los hiperparámetros previamente elegidos mediante validación cruzada temporal.
2. La semana siguiente ($W + 1$) se destina a la fase de prueba. En esta etapa se utilizan los datos reales sin filtrar, los cuales pueden incluir comportamientos anómalos o eventos asociados a cortes de hoja. El modelo genera predicciones cada tres horas a partir de la información disponible hasta ese instante, simulando un escenario de monitoreo continuo.
3. Se calcula el error entre las predicciones del modelo —entrenado bajo condiciones normales— y los valores observados. Un incremento sustancial

del error respecto al umbral normal definido se interpreta como una posible anomalía en el sensor.

Este esquema permite comparar el comportamiento reciente del sistema con su operación estable previa al evento, posibilitando una **detección temprana de anomalías**.

Además, el análisis se focaliza en los sensores que presentan la mayor conectividad hacia la serie *CH*, según los resultados del análisis de *spillovers*. Esto permite priorizar aquellos sensores cuya influencia en la ocurrencia del evento es más significativa, incrementando la capacidad de identificar **patrones precursores de corte de hoja**.

En el Anexo A2 se presentan las máscaras booleanas correspondientes a la semana previa de cada evento, junto con los intervalos seleccionados como válidos para el entrenamiento.

En el Cuadro 3.4.1 se detallan las fechas de inicio y término de los periodos de entrenamiento (semanas sin anomalías), así como los intervalos de prueba, que se extienden hasta dos bloques de tres horas posteriores al evento de corte. En esta tabla también se indica el número de ventana inmediatamente anterior al corte de hoja, dado que el objetivo es **anticipar la anomalía y determinar qué sensores presentan mayores errores de predicción**.

Finalmente, el proceso de predicción se implementó de forma incremental, generando pronósticos cada tres horas desde el lunes a las 00:00:00, extendiendo progresivamente la ventana de análisis. Este enfoque permite observar cómo evoluciona el error del modelo conforme se aproxima el momento crítico del corte y evaluar qué sensores manifiestan desviaciones anómalas con mayor anticipación temporal.

Cuadro 3.4.1: Ventanas train y test para el entrenamiento

CH	Fecha inicio train	Fecha final train	Fecha inicio test	Fecha final previa inicio CH	n° ventana previa a CH
1	2025-01-01 00:00:00	2025-01-04 23:55:00	2025-01-05 00:00:00	2025-01-06 03:00:00	9
2	2025-01-09 00:00:00	2025-01-12 23:55:00	2025-01-13 00:00:00	2025-01-15 21:00:00	23
3	2025-01-27 00:00:00	2025-01-31 23:55:00	2025-02-01 00:00:00	2025-02-09 21:00:00	71
4	2025-02-04 00:00:00	2025-02-07 23:55:00	2025-02-08 00:00:00	2025-02-11 15:00:00	29
5	2025-02-18 00:00:00	2025-02-23 23:55:00	2025-02-24 00:00:00	2025-03-01 15:00:00	45
6	2025-02-18 00:00:00	2025-02-23 23:55:00	2025-02-24 00:00:00	2025-03-02 06:00:00	50
7	2025-02-24 00:00:00	2025-02-27 23:55:00	2025-02-28 00:00:00	2025-03-03 06:00:00	12
8	2025-03-05 00:00:00	2025-03-09 23:55:00	2025-03-10 00:00:00	2025-03-12 09:00:00	19
9	2025-03-17 00:00:00	2025-03-23 23:55:00	2025-03-24 00:00:00	2025-03-28 00:00:00	32
10	2025-03-30 00:00:00	2025-03-30 23:55:00	2025-03-31 00:00:00	2025-04-03 18:00:00	30
11	2025-04-11 00:00:00	2025-04-13 23:55:00	2025-04-14 00:00:00	2025-04-15 06:00:00	10
12	2025-04-11 00:00:00	2025-04-13 23:55:00	2025-04-14 00:00:00	2025-04-20 00:00:00	48
13	2025-04-22 00:00:00	2025-04-27 23:55:00	2025-04-28 00:00:00	2025-05-02 21:00:00	39
14	2025-05-15 00:00:00	2025-05-18 23:55:00	2025-05-19 00:00:00	2025-05-23 18:00:00	38

3.4.2. Cálculo del MSE por ventana

El cálculo del **Error Cuadrático Medio** se realiza de manera independiente para cada **ventana temporal de tres horas**. Esto genera una secuencia de errores a lo largo del tiempo, permitiendo analizar la evolución del desempeño del modelo y detectar aumentos localizados del error, los cuales pueden indicar un comportamiento anómalo de los sensores o un cambio en las condiciones del proceso.

Para cada sensor y modelo m , el error en una ventana específica se calcula como:

$$MSE_{m,vent}^{CH} = \frac{1}{n} \sum_{i=1}^n (y_{i,vent}^{CH} - \hat{y}_{i,m,vent}^{CH})^2 \quad (3.4.1)$$

donde,

- $y_{i,vent}^{CH}$: valor real de la muestra i dentro de la ventana.
- $\hat{y}_{i,m,vent}^{CH}$: predicción del modelo m para la muestra i en la misma ventana.

- n : número de muestras de test en la ventana.
- m : modelo de regresión utilizado (KNN , LR , RF , SVR o XGB).
- $vent$: ventana de 3 horas donde se calcula el error.

Esta estructura permite monitorear el desempeño del modelo de manera local en el tiempo, evaluando si la predicción de cada sensor se mantiene cercana a su comportamiento normal o si se producen desviaciones significativas.

3.4.3. Monitoreo temporal del MSE y activación de umbrales de alerta

Para generar distintos niveles de alerta, se definen los factores $F \in \{1, 2, 3, 4, 5\}$, escalando el umbral máximo 3.3.2 de la siguiente manera:

$$Umbrales_{\times F}^{m,s} = Umbrales_{MSE_{m\acute{a}x}}^{m,s} \times F. \quad (3.4.2)$$

En la fase de prueba, si el error observado $MSE_{test}^{m,s}$ supera un umbral, se interpreta que el modelo ha salido de su rango de desempeño normal. Un valor elevado de F indica un comportamiento anómalo potencial, el cual puede estar asociado a condiciones operativas irregulares o a la ocurrencia de un evento de corte de hoja.

Este enfoque permite identificar de manera oportuna ventanas donde los sensores presentan un comportamiento anómalo y proporciona un marco cuantitativo para decisiones operativas, como reentrenamiento del modelo o inspecciones preventivas.

3.4.4. Indicador global de desempeño

Para evaluar cuál modelo detecta mejor anomalías asociadas a los cortes de hoja (CH), se propone un **indicador global** basado en el número promedio de sensores que superan un umbral crítico ($Umbrales_{\times 5}$) por evento:

$$g_m = \frac{1}{n} \sum_{i=1}^n w_i \quad (3.4.3)$$

donde,

- n : número total de eventos de corte de hoja considerados en el período de estudio.
- w_i : número de sensores que superaron el umbral $\times 5$ durante el evento CH_i para el modelo m .
- m : modelo de regresión evaluado (KNN , LR , RF , SVR o XGB).

En el período analizado se identificaron un total de $n = 14$ eventos de corte de hoja, los cuales se utilizan para el cálculo del indicador global de desempeño. Este indicador refleja la **sensibilidad del modelo** frente a desviaciones: a mayor g_m , mayor capacidad de detectar anomalías.

3.4.5. Selección del modelo óptimo

El modelo con el mayor valor de g_m se considera el más efectivo para la detección temprana de anomalías:

$$\text{Modelo óptimo: } \sup_m g_m$$

Esta metodología permite comparar todos los modelos bajo un criterio homogéneo: el número de sensores que superan el umbral indica la capacidad de anticipar desviaciones relevantes en el proceso.

Capítulo 4

Resultados

4.1. Conectividad

4.1.1. Ranking

El ranking de sensores utilizado en este estudio se presenta en la Tabla 4.1.1.

El análisis del ranking de conectividad hacia la serie CH revela la presencia recurrente de sensores pertenecientes a distintas etapas del proceso, lo que indica que la ocurrencia del corte de hoja no responde a una única causa local, sino a una combinación de factores distribuidos a lo largo de la línea de secado.

Destacan los sensores 3, 6 y 36, los cuales mantienen una influencia alta y constante en múltiples cortes. Su ubicación en las etapas de depuración y formación sugiere que pueden tener un efecto acumulativo que se manifiesta en los eventos de corte posteriores.

Asimismo, los sensores 75 (tensión) y 96 (velocidad de hoja) confirman relaciones causales físicamente comprobadas, dado que variaciones en la tensión o velocidad son causas directas de la rotura o pérdida de continuidad de la hoja. La alta conectividad de estos sensores hacia CH refuerza la validez del enfoque de conectividad empleado, ya que identifica como influyentes variables reconocidas por la operación de planta como precursoras reales del evento.

Por otra parte, la aparición repetida de sensores como 18, 92 y 95 en la etapa de secado sugiere que condiciones de humedad, temperatura o gradientes de presión

cercanos al final del proceso también pueden contribuir a la inestabilidad que antecede al corte.

En conjunto, estos resultados permiten distinguir zonas del proceso con influencia estructural sobre la ocurrencia de cortes de hoja, aportando una base cuantitativa para orientar estrategias de monitoreo preventivo y ajuste operativo.

La fuerte coincidencia entre los sensores influyentes de CH-5, CH-6 y CH-7 indica que estos tres cortes consecutivos probablemente fueron originados por un mismo patrón de condiciones operativas anómalas que se mantuvo durante un intervalo prolongado.

Los sensores 3, 6, 18, 36, 75 y 95 se repiten en los tres eventos con alta conectividad hacia CH, lo que sugiere que las perturbaciones en etapas tempranas del proceso (depuración, aproximación y formación) se propagaron aguas abajo, afectando la tensión y velocidad de la hoja en la zona de prensas y secado.

En conjunto, esta recurrencia apunta a un mismo modo de falla, es decir, un patrón operativo que generó tres cortes de hoja consecutivos antes de que el sistema lograra estabilizarse.

Cuadro 4.1.1: Ranking de sensores según su conectividad hacia la variable de corte de hoja (CH). Para cada columna CH_k , la columna n° representa la posición del ranking, donde $n^\circ = 1$ identifica al sensor con mayor conectividad dirigida hacia CH_k . En cada celda se indica el número del sensor, y entre paréntesis el valor de la conectividad estimada.

n°	CH-1	CH-2	CH-3	CH-4	CH-5	CH-6	CH-7	CH-8	CH-9	CH-10	CH-11	CH-12	CH-13	CH-14
1	6 (2.12)	3 (2.25)	28 (1.78)	8 (1.93)	18 (1.85)	18 (1.91)	2 (2.12)	17 (2.17)	70 (1.77)	92 (2.32)	38 (2.86)	9 (2.65)	81 (1.41)	18 (2.91)
2	96 (1.86)	17 (1.76)	2 (1.61)	3 (1.86)	28 (1.85)	3 (1.89)	6 (1.58)	67 (1.92)	36 (1.76)	19 (2.32)	2 (1.65)	57 (2.35)	66 (1.31)	1 (2.3)
3	2 (1.52)	38 (1.54)	35 (1.58)	6 (1.34)	95 (1.84)	28 (1.73)	5 (1.57)	6 (1.88)	96 (1.59)	6 (1.8)	93 (1.48)	75 (2.09)	69 (1.3)	92 (1.91)
4	1 (1.4)	70 (1.31)	26 (1.54)	22 (1.31)	3 (1.8)	95 (1.68)	9 (1.55)	3 (1.88)	6 (1.53)	84 (1.78)	47 (1.42)	26 (1.98)	18 (1.29)	3 (1.59)
5	47 (1.32)	32 (1.27)	19 (1.47)	12 (1.28)	36 (1.77)	36 (1.61)	67 (1.47)	69 (1.53)	75 (1.53)	83 (1.77)	45 (1.37)	63 (1.88)	15 (1.28)	15 (1.48)
6	75 (1.17)	30 (1.26)	3 (1.45)	11 (1.26)	10 (1.7)	38 (1.6)	36 (1.45)	18 (1.48)	3 (1.33)	36 (1.39)	92 (1.35)	62 (1.87)	19 (1.21)	56 (1.4)
7	81 (1.13)	95 (1.26)	24 (1.45)	54 (1.25)	38 (1.6)	23 (1.49)	1 (1.32)	36 (1.42)	56 (1.32)	30 (1.3)	17 (1.31)	40 (1.81)	71 (1.2)	69 (1.36)
8	28 (1.08)	50 (1.17)	36 (1.38)	61 (1.15)	8 (1.54)	96 (1.44)	10 (1.3)	25 (1.36)	53 (1.28)	2 (1.2)	6 (1.3)	39 (1.78)	4 (1.19)	49 (1.24)
9	56 (1.07)	27 (1.16)	66 (1.38)	95 (1.1)	23 (1.49)	25 (1.43)	20 (1.25)	61 (1.3)	28 (1.25)	47 (1.18)	68 (1.3)	82 (1.78)	96 (1.19)	12 (1.2)
10	95 (1.06)	47 (1.16)	8 (1.33)	17 (1.05)	25 (1.43)	15 (1.4)	12 (1.15)	66 (1.25)	2 (1.19)	59 (1.09)	95 (1.3)	41 (1.76)	24 (1.14)	5 (1.1)
11	3 (1.05)	2 (1.02)	27 (1.27)	80 (1.04)	47 (1.37)	6 (1.35)	47 (1.15)	75 (1.25)	59 (1.16)	59 (1.09)	23 (1.29)	24 (1.73)	27 (1.12)	24 (1.08)
12	57 (1.04)	37 (1.02)	6 (1.26)	10 (1.01)	6 (1.34)	47 (1.35)	17 (1.13)	62 (1.18)	26 (1.12)	7 (1.07)	15 (1.26)	59 (1.61)	32 (1.11)	52 (1.08)
13	83 (1.04)	33 (0.98)	17 (1.25)	5 (1)	81 (1.17)	94 (1.2)	85 (1.12)	96 (1.16)	27 (1.12)	15 (1.94)	28 (1.23)	4 (1.58)	16 (1.1)	28 (1.07)
14	84 (1.02)	1 (0.93)	14 (1.24)	87 (0.94)	94 (1.17)	26 (1.14)	75 (1.1)	20 (1.15)	69 (1.1)	50 (1.02)	66 (1.23)	36 (1.56)	36 (1.1)	75 (1.05)
15	66 (0.99)	36 (0.91)	75 (1.24)	33 (0.92)	77 (1.09)	77 (1.14)	78 (1.09)	71 (1.11)	43 (1.07)	75 (1.01)	25 (1.06)	71 (1.55)	68 (1.09)	96 (1.05)

4.1.2. Correlación

El análisis de correlación de Pearson realizado sobre los sensores muestra que los pares de sensores con coeficiente superior a 0.9 se concentran principalmente dentro de la misma etapa del proceso de secado. Esto indica que los sensores que se encuentran físicamente cercanos o que miden variables relacionadas funcionalmente tienden a presentar un comportamiento altamente similar, reflejando redundancia en la información registrada.

En el Cuadro 4.1.2 se observa que varios pares de sensores altamente correlacionados fueron eliminados siguiendo el criterio de prioridad por etapa. Además, los sensores conservados corresponden principalmente a los que están más cercanos al inicio de la máquina, ya que los sensores ubicados más atrás pueden reflejar efectos de rebote o retrasos en el proceso, y por tanto no representan información independiente. Los sensores con varianza nula también fueron descartados al no aportar información significativa. Asegurando que los modelos de predicción y detección de anomalías se entrenen con datos informativos y no redundantes. Este filtrado previo permite reducir la dimensionalidad del conjunto de datos sin perder representatividad del comportamiento del proceso, facilitando la identificación de patrones y relaciones críticas entre variables.

Cuadro 4.1.2: Correlación Pearson >0.9

CH	Pares de sensores con correlación >0.9	Sensores eliminados	Sensores con varianza = 0	Cantidad de sensores que se mantienen
1	(1,2), (83,84)	2,84	47	12
2	(1, 2), (36, 37)	2, 37	47	12
3	0	0	0	15
4	0	0	0	15
5	(23,25)	25	47	13
6	(23,25)	25	47	13
7	(1,2), (17,20)	2, 20	47	12
8	(17,20)	20	0	14
9	0	0	0	15
10	(83,84)	84	47	13
11	(23,25)	25	47	13
12	(40,41), (40,62), (40,63), (41,62), (71,63), (57,59), (62,63)	41, 59, 62, 63	0	11
13	(68,69)	69	0	14
14	(1,15)	15	0	14

4.2. Desempeño de modelos de entrenamiento y selección de hiperparámetros

4.2.1. Umbrales de detección

En los cuadros correspondientes 4.2.1, 4.2.2, 4.2.3 4.2.4 y 4.2.5 se presentan los resultados obtenidos para cada modelo de predicción y evento de corte de hoja (CH). Para cada sensor se indica si el valor de MSE_{test} superó el umbral de normalidad multiplicado por cinco ($F=5$), junto con la ventana temporal en la que ocurrió dicha superación. Este análisis permite identificar los sensores que presentan desviaciones más tempranas y, por lo tanto, podrían actuar como indicadores precursores de los eventos de corte de hoja.

En términos generales, se observa que algunos sensores muestran de manera consistente una superación del umbral en múltiples ventanas previas al evento, mientras que otros lo hacen de forma esporádica. La información registrada en las tablas permite comparar el desempeño de los distintos modelos (KNN, LR, RF, SVR y XGB) y su sensibilidad a desviaciones en los sensores.

De manera más específica, se pueden destacar los siguientes patrones:

- **Sensores consistentes:** Algunos sensores, como los ubicados en los grupos 3 y 4 (aproximación y formación), presentan superaciones recurrentes del umbral en casi todos los modelos, lo que indica que podrían estar altamente correlacionados con la ocurrencia de cortes de hoja. Por ejemplo, los sensores 2, 3, 6, 8, 14, 17 y 36 aparecen de forma consistente entre KNN, LR, RF, SVR y XGB.
- **Variabilidad entre modelos:** Si bien los sensores mencionados muestran consistencia, otros sensores aparecen solo en algunos modelos o en eventos específicos. Esto refleja que la capacidad de detección temprana depende del tipo de modelo utilizado y de la sensibilidad a diferentes patrones de error.
- **Eventos sin detección temprana:** Algunos eventos de CH (por ejemplo, los identificados como 7, 8 y 9) presentan pocos o ningún sensor que supere el umbral de manera significativa. Esto sugiere que dichos eventos pueden no ser predecibles mediante la información disponible en los sensores, o bien que los desvíos asociados ocurren muy próximos al instante del corte, limitando

544.2. Desempeño de modelos de entrenamiento y selección de hiperparámetros

la capacidad de detección temprana de los modelos.

- Superación múltiple de ventanas:** La cantidad de ventanas en las que un sensor supera el umbral (indicada entre paréntesis en las tablas) refleja la persistencia del comportamiento anómalo. Sensores con un número elevado de ventanas superadas podrían ser considerados como alertas tempranas confiables. Por ejemplo, para el evento 3, varios sensores superan el umbral en un rango amplio de ventanas (ventanas 14–71), indicando una anomalía prolongada antes del corte.

Cuadro 4.2.1: Sensores que superaron el umbral $\times 5$ para el modelo *KNN*.
() es el n^o de ventanas en que fue superado el umbral $\times 5$.

KNN	CH	n ^o sensor >4
	CH-1	6, 57 (2-4), (2-9)
	CH-2	1, 27, 38 (7-8), (7-8), (18-26)
	CH-3	2, 3, 6, 8, 14, 17, 19, 24, 26, 27, 28, 35, 36, 66, 75 (15-71), (14-71), (30-71), ((18-53)-(65-71)), (14-71), (15-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (21-22), (14-71), (14-71), (14-71)
	CH-4	3, 5, 6, 8, 11, 12, 17, 22, 54, 61 (15-29), (19-29), (15-29), (19-29), (20-29), (20-29), (20-29), (20-29), (19-29), (19-29)
	CH-5	36, 77 (14-45), (14-35)
	CH-6	36, 38, 77, 96 (14-50), (47-50), (15-16), (47-50)
	CH-7	-
	CH-8	-
	CH-9	-
	CH-10	6, 7, 8, 15, 30, 59, 83, 92 (5-30), (10-30), (9-30), (15-30), (6-30), (20-30), (13-30), (13-30), (21-30)
	CH-11	15, 17, 23, 28, 38, 45, 66, 68, 92, 93 (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10)
	CH-12	24, 28, 36, 39, 40, 57, 82, 71, 75 (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48)
	CH-13	68 (6-39)
	CH-14	-

4.2. Desempeño de modelos de entrenamiento y selección de hiperparámetros 55

Cuadro 4.2.2: Sensores que superaron el umbral $\times 5$ para el modelo *LR*.
() es el n^o de ventanas en que fue superado el umbral $\times 5$.

LR	CH	n ^o sensor >4
	CH-1	-
	CH-2	1 (12-23)
	CH-3	2, 3, 6, 8, 14, 17, 19, 24, 26, 27, 28, 35, 36, 66, 75 (15-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (15-71), (14-71), (14-71), (14-71)
	CH-4	3, 5, 6, 8, 10, 11, 12, 17, 22, 33, 54, 61, 80, 87, 95 (15-29), (19-29), (15-29), (19-29), (20-29), (17-29), (19-29), (19-29), (27-29), (19-29), (19-29), (19-29), (19-29), (23-29)
	CH-5	36, 77 (14-45), (15-17)
	CH-6	23, 26, 28, 36, 38, 77, 96 (47-50), (47-50), (46-50), (14-50), (47-50), ((14-20), (23-26), (47-50)), (47-50)
	CH-7	-
	CH-8	-
	CH-9	-
	CH-10	6, 7, 8, 15, 83, 92 (10-30), (10-30), (8-30), (13-30), (12-30), (21-30)
	CH-11	2, 6, 15, 17, 23, 28, 38, 45, 66, 68, 92, 93, 95 (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10)
	CH-12	4, 9, 24, 26, 36, 39, 40, 57, 71, 75, 82 (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48)
	CH-13	19, 24, 32, 36, 66, 68, 71 (12-39), (11-39), (36-39), (12-39), (12-39), (6-39), (6-39)
	CH-14	24, 52 (23-38), (38)

564.2. Desempeño de modelos de entrenamiento y selección de hiperparámetros

Cuadro 4.2.3: Sensores que superaron el umbral $\times 5$ para el modelo *RF*.

() es el n^o de ventanas en que fue superado el umbral $\times 5$.

RF	CH	n ^o sensor >4
	CH-1	57 (3)
	CH-2	38,1 (5-23), (12-23)
	CH-3	2,3,6,8,14,17,19,24,26,27,28,35,36,66,75 (15-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (15-71), (14-71), (14-71), (14-71), (14-71)
	CH-4	3,5,6,8,11,17,22,54,61,87 (15-29), (19-29), (15-29), (17-29), (26-29), (20-29), (19-29), (19-29), (19-29), (19-29), (25-29)
	CH-5	36,77 (14-45), (14-40)
	CH-6	15,28,36,38,96 (26-50), (26-50), (14-50), (47-50), (46-50)
	CH-7	-
	CH-8	-
	CH-9	-
	CH-10	6,7,8,15,30,59,83,92 (4-30), (9-30), (9-30), (15-30), (4-30), (21-30), (12-30), (21-30)
	CH-11	2,15,17,23,28,38,45,66,68,92,93 ((2), (7), (10), (12-14)), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10)
	CH-12	24,28,36,39,40,57,71,75,82 (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48)
	CH-13	24,68 (24-39), (6-39)
	CH-14	-

4.2. Desempeño de modelos de entrenamiento y selección de hiperparámetros 57

Cuadro 4.2.4: Sensores que superaron el umbral $\times 5$ para el modelo *SVR*.
() es el n^o de ventanas en que fue superado el umbral $\times 5$.

SVR	CH	n ^o sensor >4
	CH-1	6,57,81 (2-6), (2-9), (2-5)
	CH-2	38,1 (9-23), (12-23)
	CH-3	2,3,6,8,14,17,19,24,26,27,28,35,36,66,75 (15-71), (20-71), (16-71), (14-71), (15-71), (14-71), (14-71), (14-71), (14-71), (14-71), (20-27), (14-71), (14-71), (14-71)
	CH-4	3,5,6,8,12,17,22,54,61 (15-29), (19-29), (15-29), (19-29), (26-29), (20-29), (19-29), (19-29), (19-29)
	CH-5	28,36,77 (46-48), (14-45), (14-36)
	CH-6	28,36,38,96 (46-50), (14-50), (47-50)
	CH-7	-
	CH-8	-
	CH-9	-
	CH-10	6,7,8,15,59,83,92 (4-30), (11-30), (9-30), (14-30), (22-30), (12-30), (21-30)
	CH-11	2,15,17,23,28,38,45,66,68,92,93 (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10)
	CH-12	24,26,36,39,40,57,71,75,82 (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48)
	CH-13	68 (6-39)
	CH-14	-

584.2. Desempeño de modelos de entrenamiento y selección de hiperparámetros

Cuadro 4.2.5: Sensores que superaron el umbral $\times 5$ para el modelo *XGB*.

() es el n° de ventanas en que fue superado el umbral $\times 5$.

XGB	CH	n° sensor >4
	CH-1	57 (2-9)
	CH-2	38,1 (5-23), (12-23)
	CH-3	2,3,8,14,17,19,24,26,27,28,35,36,66,75 (15-71), (14-71), (16-46), (14-71), (15-71), (14-71), (14-71), (14-71), (14-71), (14-71), (14-71), (18-30), (14-71), (14-71), (14-71)
	CH-4	3,5,6,8,11,12,17,22,54,61,87 (15-29), (19-29), (15-29), (19-29), (26-29), (9-17), (26-29), (20-29), (19-29), (19-29), (19-29), (25-29)
	CH-5	36,77 (14-45), (14-36)
	CH-6	15,28,36,38,96 (26-50), (28-50), (14-50), (47-50), (47-50), (46-50)
	CH-7	-
	CH-8	-
	CH-9	-
	CH-10	6,7,8,15,30,59,83,92 (4-30), (9-30), (9-30), (15-30), (5-30), (21-30), (12-30), (21-30)
	CH-11	2,15,17,23,28,38,45,66,68,92,93 (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10), (1-10)
	CH-12	24,26,36,39,40,57,71,75,82 (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48), (1-48)
	CH-13	24,68 (24-39), (6-39)
	CH-14	-

De igual forma, en el Anexo A1 se presenta la gráfica que muestra la evolución temporal del nivel de umbral superado en cada ventana de análisis. En este gráfico, se representa con un número del 1 al 5 el factor de umbral (F) excedido en cada intervalo, de acuerdo con la magnitud del error *MSE* observado. Este procedimiento se repite para cada evento de corte de hoja y para cada modelo evaluado, permitiendo visualizar de manera comparativa cómo evoluciona el nivel de desviación del error respecto al umbral de normalidad a lo largo del tiempo previo al evento anómalo.

A modo de ejemplo, en la Figura 4.2.1 se muestra la evolución del error del modelo KNN para el evento de corte de hoja 2, indicando en qué ventanas de tiempo el modelo supera los distintos niveles del umbral de normalidad ($\times 1$, $\times 2$, $\times 3$, $\times 4$, $\times 5$). El gráfico evidencia cómo el error del modelo aumenta progresivamente a medida que se aproxima el evento, reflejando una degradación en la capacidad predictiva frente a condiciones fuera de lo normal. La superación de umbrales más altos

4.2. Desempeño de modelos de entrenamiento y selección de hiperparámetros 59

sugiere una desviación significativa respecto al comportamiento esperado del sensor, interpretándose como una señal temprana de anomalía en el proceso.

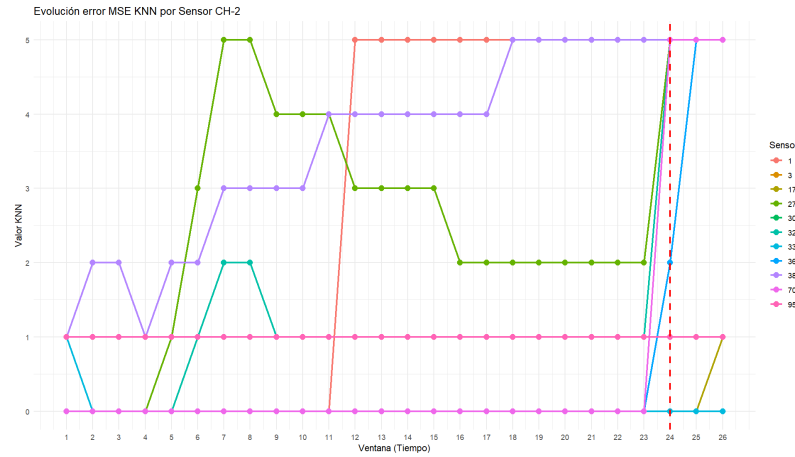


Figura 4.2.1: Evolución temporal del error del modelo KNN para el CH-2, indicando los niveles de umbral superados.

4.2.2. Conclusiones preliminares de las tablas de detección de umbrales

1. Los sensores de los grupos de aproximación y formación (3 y 4) muestran la mayor consistencia como precursores de cortes de hoja, siendo detectables por todos los modelos evaluados.
2. Sensores específicos como 2, 3, 6, 8, 14, 17, 36 y 75 son recurrentes en múltiples eventos y modelos, lo que sugiere que podrían tener un papel clave en la anticipación de los cortes.
3. Modelos como Linear Regression y Random Forest presentan patrones similares de detección a KNN, mientras que SVR y XGB muestran ligeras variaciones en los sensores detectados, indicando diferencias al calcular el error.
4. La persistencia de la superación del umbral (cantidad de ventanas afectadas) es un indicador importante de relevancia de un sensor como precursor: sensores con un rango amplio de ventanas superadas podrían proporcionar alertas tempranas más confiables.

5. Algunos eventos de corte no presentan detección temprana consistente, lo que indica que no todos los cortes son igualmente previsibles, posiblemente debido a variabilidad en la operación del proceso o a factores externos no medidos.

4.3. Identificación del modelo óptimo

Para evaluar la capacidad de cada modelo en la detección temprana de anomalías asociadas a los eventos de corte de hoja (CH), se calcularon los indicadores $w_i/14$ y g_m correspondientes a cada modelo evaluado (KNN, LR, RF, SVR y XGB).

En la Tabla 4.3.1 se muestran los valores de $w_i/14$ para cada evento de corte de hoja y cada modelo. Estos valores representan la proporción de sensores que superaron el umbral de error en cada CH, normalizados respecto al total de sensores considerados.

Cuadro 4.3.1: Proporción de sensores que superaron el umbral $F = 5$ para cada modelo y evento de corte de hoja.

w_i	KNN	LR	RF	SVR	XGB
CH 1	$\frac{2}{12}$	$\frac{0}{12}$	$\frac{1}{12}$	$\frac{3}{12}$	$\frac{1}{12}$
CH 2	$\frac{3}{12}$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{2}{12}$	$\frac{2}{12}$
CH 3	$\frac{15}{15}$	$\frac{15}{15}$	$\frac{15}{15}$	$\frac{15}{15}$	$\frac{14}{15}$
CH 4	$\frac{10}{15}$	$\frac{15}{15}$	$\frac{10}{15}$	$\frac{9}{15}$	$\frac{11}{15}$
CH 5	$\frac{2}{13}$	$\frac{2}{13}$	$\frac{2}{13}$	$\frac{3}{13}$	$\frac{2}{13}$
CH 6	$\frac{4}{13}$	$\frac{7}{13}$	$\frac{5}{13}$	$\frac{4}{13}$	$\frac{5}{13}$
CH 7	$\frac{0}{12}$	$\frac{0}{12}$	$\frac{0}{12}$	$\frac{0}{13}$	$\frac{0}{12}$
CH 8	$\frac{0}{14}$	$\frac{0}{14}$	$\frac{0}{14}$	$\frac{0}{14}$	$\frac{0}{14}$
CH 9	$\frac{0}{15}$	$\frac{0}{15}$	$\frac{0}{15}$	$\frac{0}{15}$	$\frac{0}{15}$
CH 10	$\frac{8}{13}$	$\frac{6}{13}$	$\frac{8}{13}$	$\frac{7}{13}$	$\frac{8}{13}$
CH 11	$\frac{10}{13}$	$\frac{13}{13}$	$\frac{11}{13}$	$\frac{11}{13}$	$\frac{11}{13}$
CH 12	$\frac{9}{11}$	$\frac{11}{11}$	$\frac{9}{11}$	$\frac{9}{11}$	$\frac{9}{11}$
CH 13	$\frac{1}{14}$	$\frac{7}{14}$	$\frac{2}{14}$	$\frac{1}{14}$	$\frac{2}{14}$
CH 14	$\frac{0}{14}$	$\frac{2}{14}$	$\frac{0}{14}$	$\frac{0}{14}$	$\frac{0}{14}$

Se observa que, para el evento CH-3, la mayoría de los sensores supera el umbral $\times 5$, lo que indica una alteración significativa y simultánea del comportamiento esperado de los sensores del sistema, tal como se muestra en la Figura correspondiente.

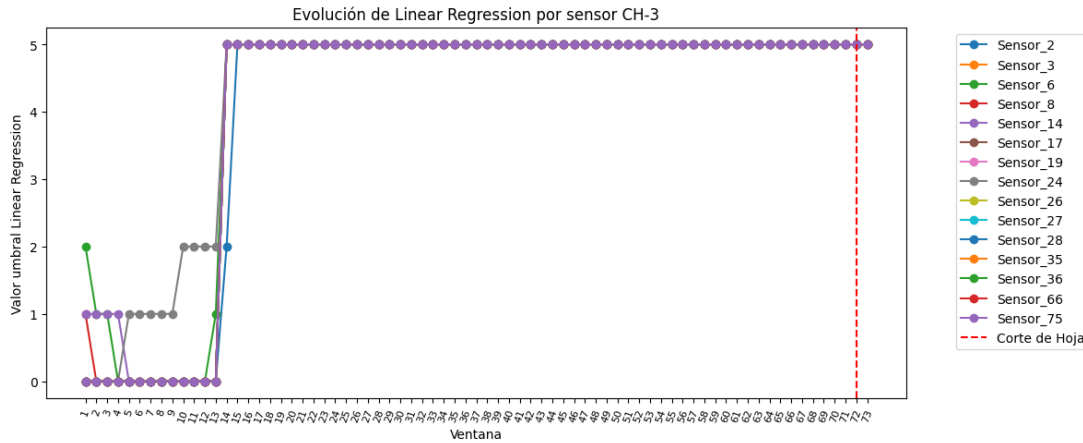


Figura 4.3.1: Evolución temporal del error del modelo LR para el CH-3, indicando los niveles de umbral superados.

Asimismo, en la Tabla 4.3.2 se presentan los valores de g_m , que resumen el desempeño global de cada modelo considerando todos los eventos de corte de hoja analizados.

Cuadro 4.3.2: Valor global g_m de cada modelo.

	KNN	LR	RF	SVR	XGB
	0.344	0.420	0.348	0.345	0.348

Como se observa, el modelo LR (Regresión Lineal) presenta el valor más alto de g_m , lo que indica que, en promedio, tiene una mayor capacidad para detectar desviaciones anómalas entre los sensores a lo largo de los distintos eventos de corte de hoja. Por esta razón, se selecciona LR como el modelo de mejor desempeño global bajo las métricas y condiciones definidas en este estudio.

Este resultado puede explicarse por varias razones inherentes al modelo de regresión lineal:

- **Estabilidad y generalización:** La regresión lineal captura de manera efectiva la tendencia central del comportamiento esperado de cada sensor. Esto permite que cualquier desviación significativa, como la generada por

un evento CH, produzca un aumento claro en el error de reconstrucción (MSE).

- **Filtrado de ruido:** La menor complejidad del modelo actúa como un filtro frente a fluctuaciones aleatorias de alta frecuencia, enfocándose en cambios estructurales en la señal. Esto es particularmente útil en procesos industriales donde las series presentan alta autocorrelación y variabilidad inherente.
- **Interpretabilidad:** A diferencia de modelos más complejos como RF, XGB o SVR, LR permite interpretar fácilmente qué sensores contribuyen a las desviaciones, facilitando la identificación de precursores de eventos y la toma de decisiones operativas.

En conjunto, estos resultados sugieren que, para la detección temprana de anomalías en entornos industriales complejos, modelos lineales bien calibrados pueden ofrecer un equilibrio óptimo entre sensibilidad, estabilidad e interpretabilidad, siendo adecuados para anticipar eventos críticos como los cortes de hoja.

Capítulo 5

Conclusión

5.1. Conclusión

El presente estudio permitió identificar y caracterizar los sensores precursores de los eventos de corte de hoja (CH) en la secadora FLAKT de CMPC, a través del análisis de conectividad, correlación y detección temprana mediante modelos de predicción basados en Machine Learning. Los resultados obtenidos aportan evidencia cuantitativa sobre los puntos críticos del proceso y las relaciones operativas que influyen en la ocurrencia de cortes:

- El análisis de ranking de conectividad hacia CH reveló que los eventos no son consecuencia de una única causa local, sino de la interacción de múltiples sensores distribuidos a lo largo de la línea de secado. Sensores como 3, 6 y 36 (ubicados en etapas de depuración y formación) muestran una influencia constante en varios cortes, sugiriendo que perturbaciones iniciales en las mediciones de los sensores se propagan a etapas posteriores.
- De manera complementaria, sensores como 75 (tensión) y 96 (velocidad de hoja) confirman relaciones causalmente consistentes con la física del proceso, reforzando la validez del enfoque de conectividad empleado.
- La repetición de sensores influyentes en cortes consecutivos (CH-5, CH-6, CH-7) indica la presencia de un mismo patrón operativo que condujo a múltiples eventos antes de la estabilización del sistema. Esta recurrencia evidencia que ciertas condiciones anómalas del proceso se mantienen a lo

largo del tiempo, afectando de manera estructural la ocurrencia de los cortes.

- En la detección de umbrales mediante modelos de Machine Learning (KNN, LR, RF, SVR y XGB), se identificaron sensores precoces y consistentes como los ubicados en los grupos de aproximación y formación (3 y 4), que superan recurrentemente los umbrales de error en múltiples ventanas temporales antes del evento. Esto demuestra su capacidad como indicadores tempranos de anomalía. La persistencia de la superación del umbral se mostró como un indicador relevante: sensores con un alto número de ventanas superadas proporcionan alertas tempranas confiables, mientras que eventos sin detección temprana reflejan la dificultad de anticipar ciertos cortes, posiblemente por factores externos no medidos o por la ocurrencia muy cercana al evento.

En conjunto, los hallazgos permiten identificar zonas críticas del proceso y sensores clave que pueden servir como precursores de cortes de hoja. Esta información proporciona una base sólida para implementar estrategias de monitoreo preventivo, ajuste operativo y mantenimiento predictivo, contribuyendo a la reducción de eventos no deseados y a la mejora de la estabilidad y eficiencia del proceso de secado.

En conclusión, el enfoque combinado de análisis de conectividad, correlación y modelos de detección de anomalías demuestra ser efectivo para anticipar eventos de corte de hoja, destacando la importancia de considerar tanto la interacción entre etapas del proceso como la persistencia de señales anómalas para una detección temprana confiable.

Bibliografía

- Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, 46(3):175–185.
- Barigozzi, M. and Brownlees, C. (2019). Nets: Network estimation for time series. *Journal of Applied Econometrics*, 34(3):347–364.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. Wiley, 5 edition.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.
- Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794.
- ciencia y tecnología, U. (2021). Predicción del corte de la hoja de celulosa mediante el uso de machine learning. Recuperado de ResearchGate.
- CMPC (2024). Cmpc es elegida por segundo año consecutivo como la empresa forestal y papelera más sostenible del mundo. Recuperado de CMPC.
- CMPC (2025). Área celulosa. Recuperado de CMPC.
- de Relaciones Económicas Internacionales, S. (2024). Informe comercial: valor de las exportaciones de servicios superan por primera vez los us\$ 2.500 millones y envíos de bienes crecen 4,7% en 2024. Recuperado de SUBREI.
- Demirer, M., Diebold, F., Liu, L., and Yilmaz, K. (2018). Estimating global bank network connectedness. *Journal of Applied Econometrics*, 33(1):1–15.
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427–431.
- Diebold, F. X. and Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*.
- Emma, L. (2025). Ai-driven predictive maintenance for smart manufacturing and industry 4.0. Recuperado de ResearchGate.

- Ferrari, D., Ravazzolo, F., and Vespignani, J. (2021). Forecasting energy commodity prices: A large global dataset sparse approach. *Energy Economics*, 98:105268.
- Forestal, I. (2025). Exportaciones forestales crecen 4,9 % en el primer trimestre impulsadas por la demanda de pulpa por parte de china. Recuperado de INFOR.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer, New York.
- Kim, J., Kim, H., Kim, H., Lee, D., and Yoon, S. (2025). A comprehensive survey of deep learning for time series forecasting: architectural diversity and open challenges. *Artificial Intelligence Review*, 58(7):1–95.
- Koop, G., Pesaran, M., and Potter, S. (1996). Impulse response analysis in non-linear multivariate models. *Journal of Econometrics*, 74(1):119–147.
- Liaw, A. and Wiener, M. (2002). Classification and regression by randomforest. *R News*, 2(3):18–22.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.
- McCowan, M. (2024). *A Mixed-Method Program Evaluation of the Advancement of the Stem Initiative in a Virginia School District with a Focus on Elementary Schools*. PhD thesis, Concordia University Texas.
- Montgomery, D. C., Peck, E. A., and Vining, G. G. (2021). *Introduction to Linear Regression Analysis*. Wiley, 6 edition.
- Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58:240–242.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12:2825–2830.
- Pesaran, M. and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters*, 58(1):17–29.
- Said, S. E. and Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607.
- scikit-learn developers (2024). Timeseriessplit – scikit-learn documentation. Accedido en 2025.
- Smola, A. J. and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222.
- Trapletti, A. and Hornik, K. (2023). *tseries: Time Series Analysis and Computational Finance*. R package version 0.10-55.

Apéndice A

Resultados extras

A1. Monitoreo temporal MSE

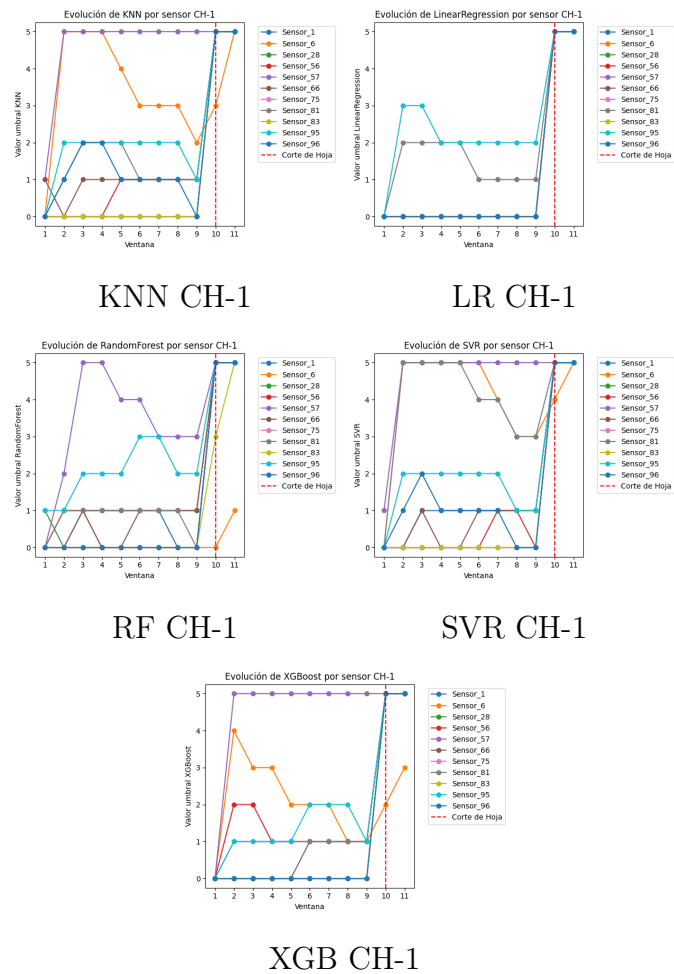
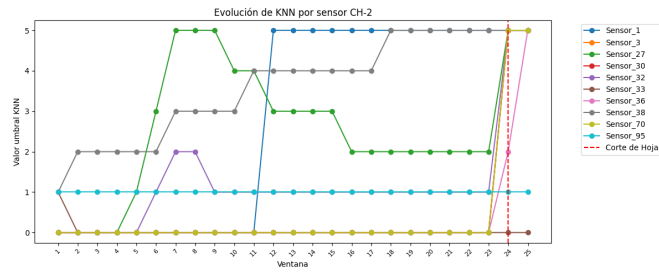
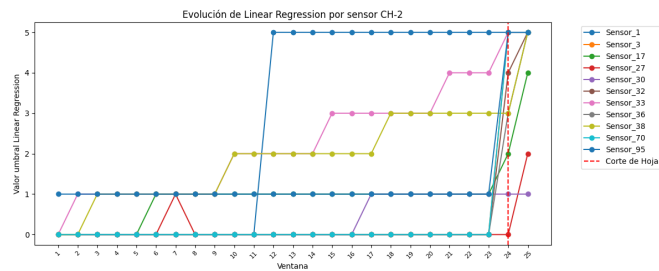


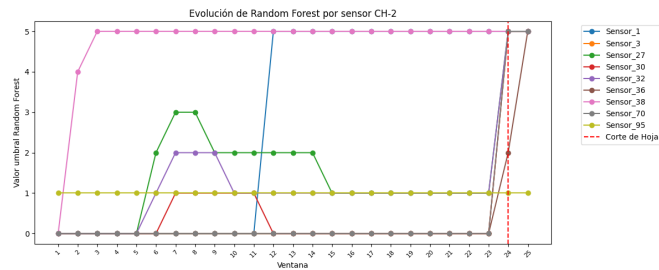
Figura A1.1: Comparación métodos CH-1



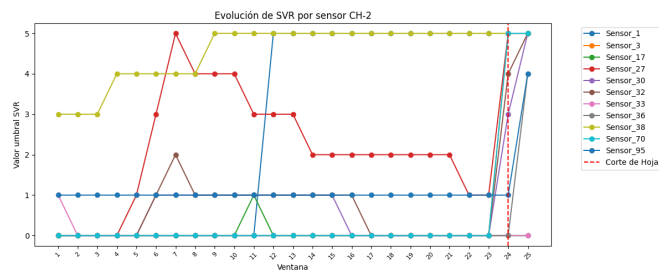
KNN CH-2



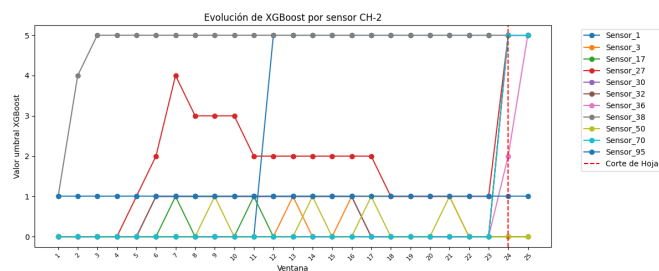
LR CH-2



RF CH-2

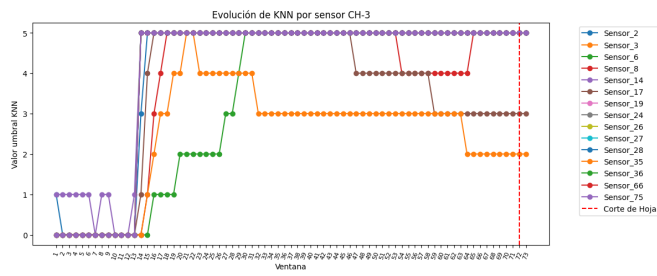


SVR CH-2

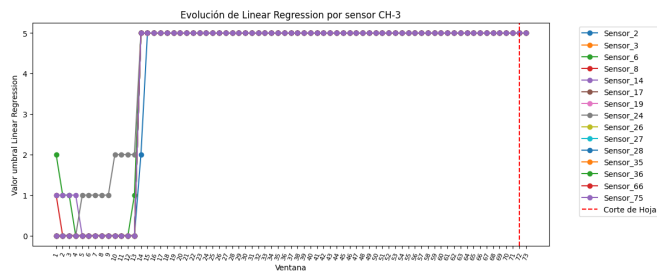


XGB CH-2

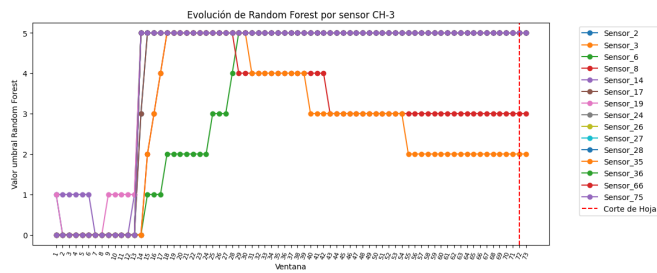
Figura A1.2: Comparación vertical de métodos en CH-2



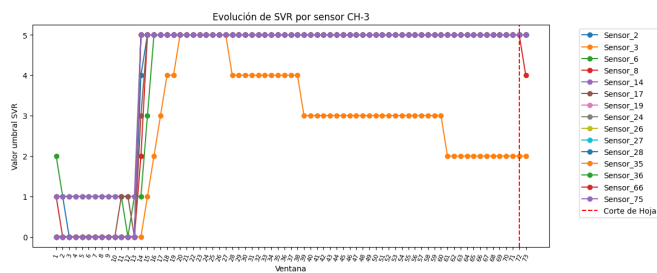
KNN CH-3



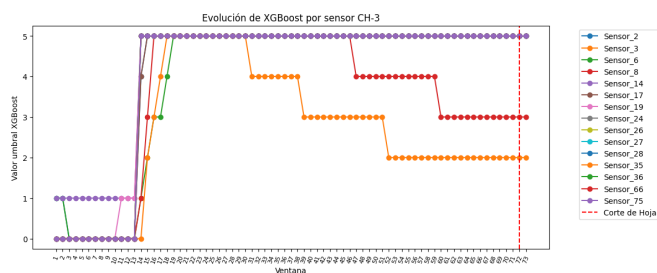
LR CH-3



RF CH-3

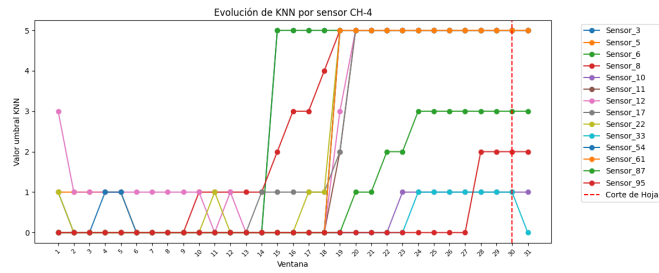


SVR CH-3

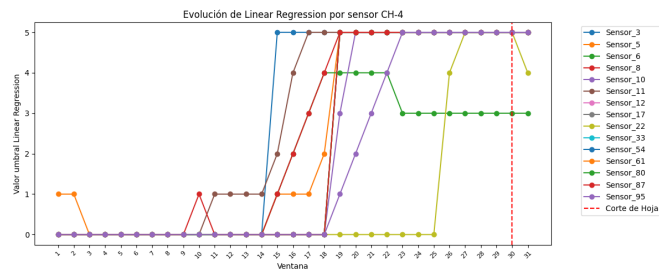


XGB CH-3

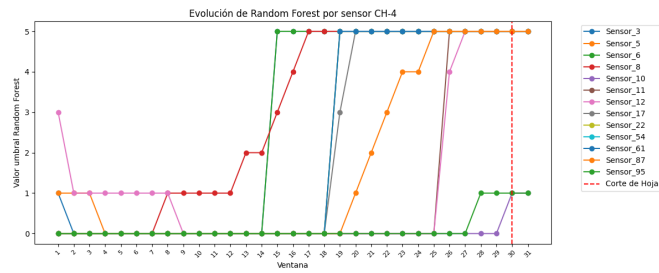
Figura A1.3: Comparación vertical de métodos en CH-3



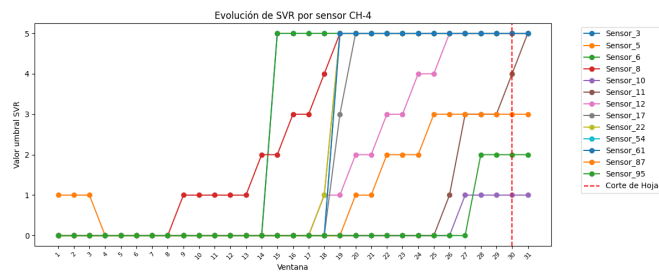
KNN CH-4



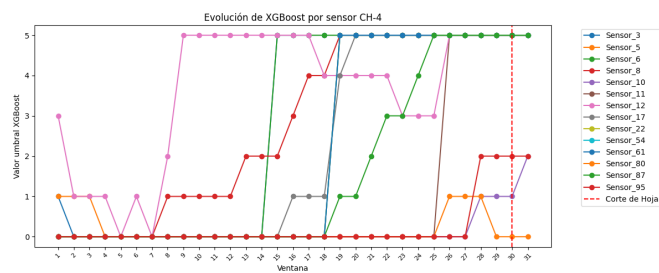
LR CH-4



RF CH-4

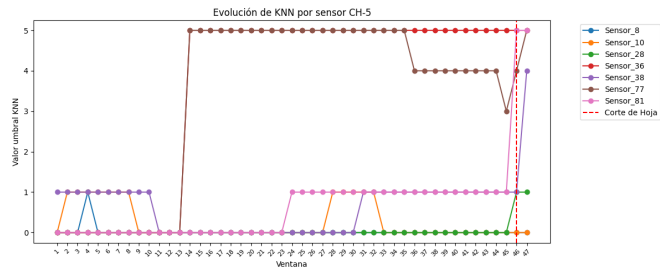


SVR CH-4

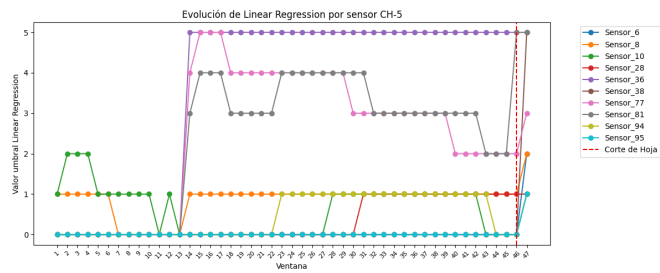


XGB CH-4

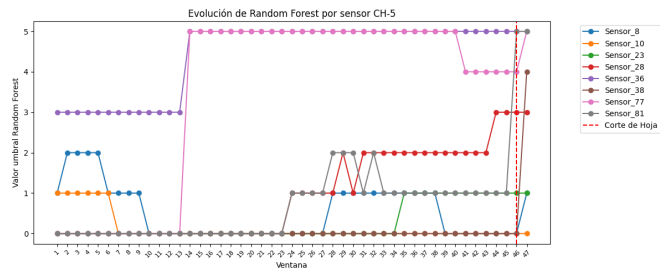
Figura A1.4: Comparación vertical de métodos en CH-4



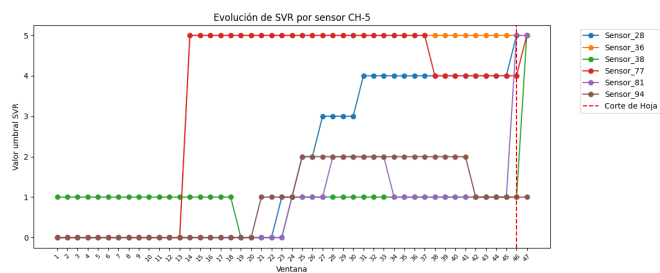
KNN CH-5



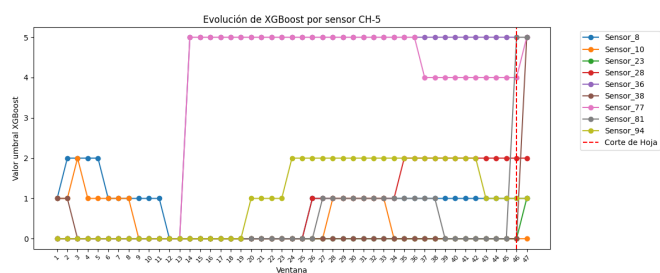
LR CH-5



RF CH-5

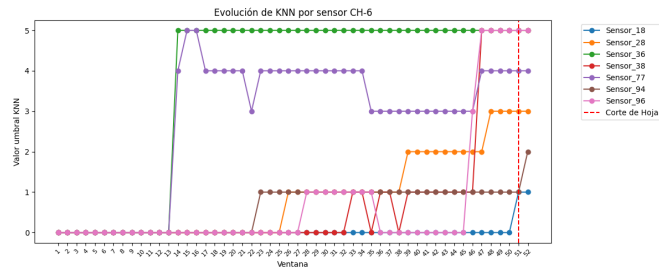


SVR CH-5

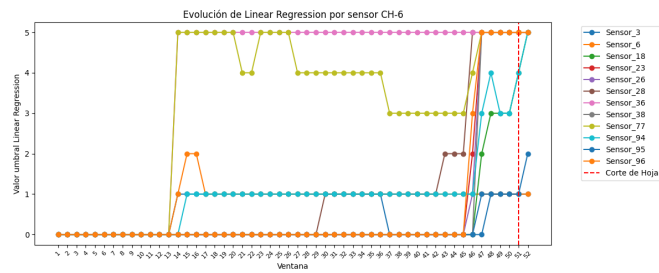


XGB CH-5

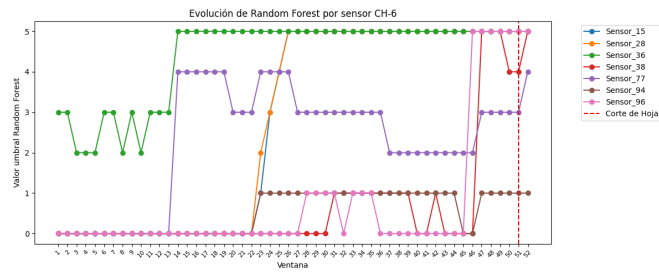
Figura A1.5: Comparación vertical de métodos en CH-5



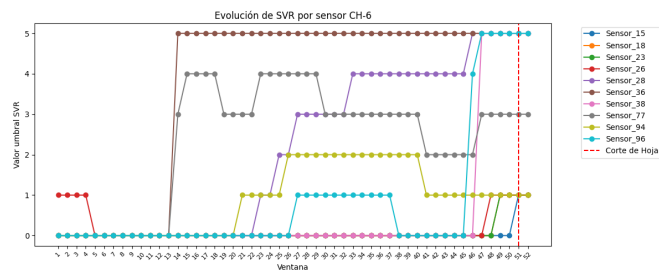
KNN CH-6



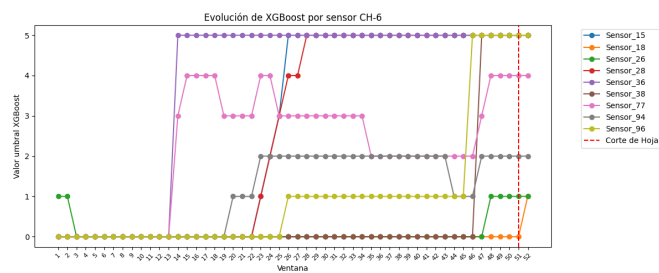
LR CH-6



RF CH-6

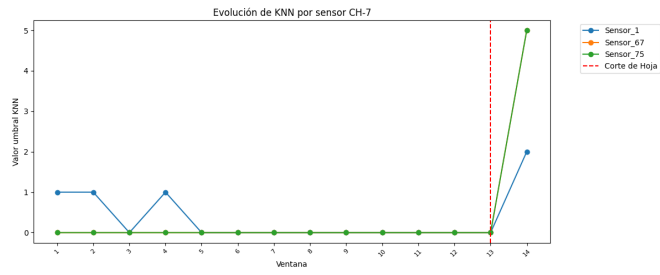


SVR CH-6

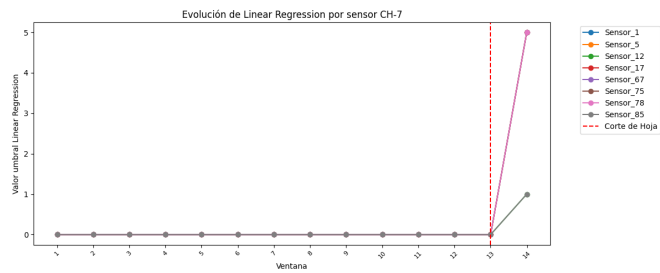


XGB CH-6

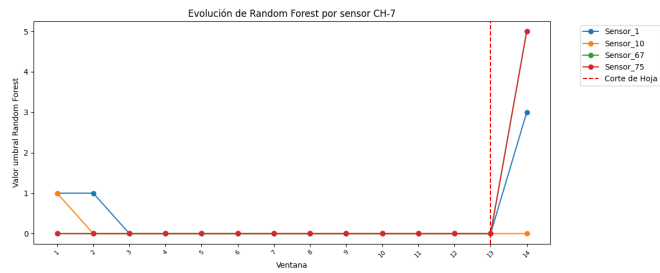
Figura A1.6: Comparación vertical de métodos en CH-6



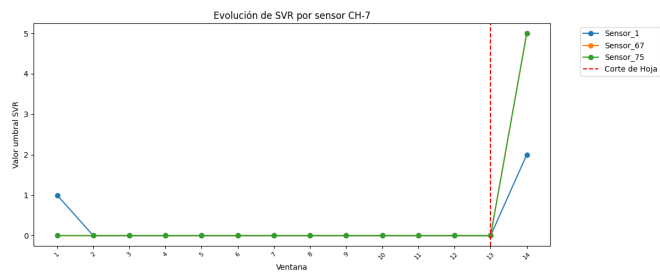
KNN CH-7



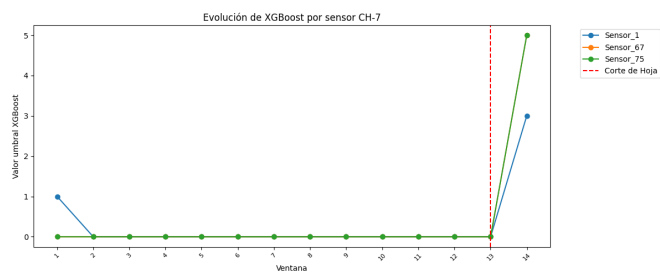
LR CH-7



RF CH-7

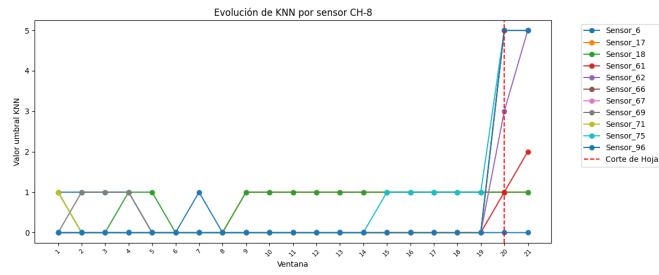


SVR CH-7

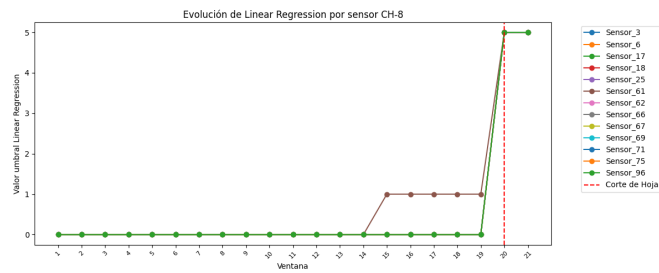


XGB CH-7

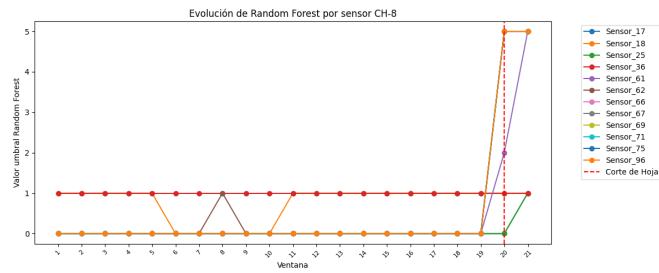
Figura A1.7: Comparación vertical de métodos en CH-7



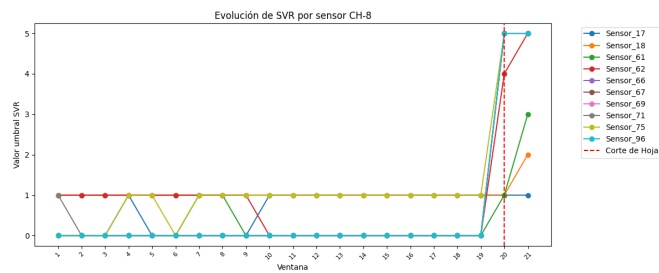
KNN CH-8



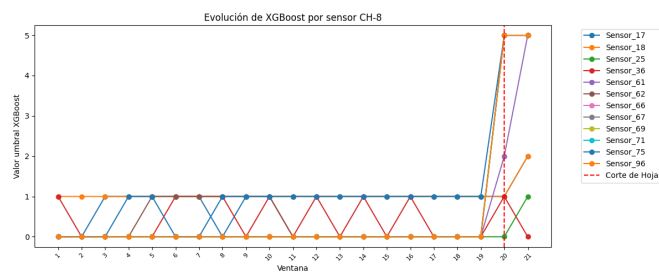
LR CH-8



RF CH-8

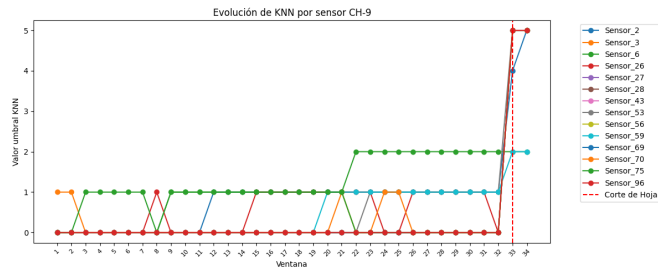


SVR CH-8

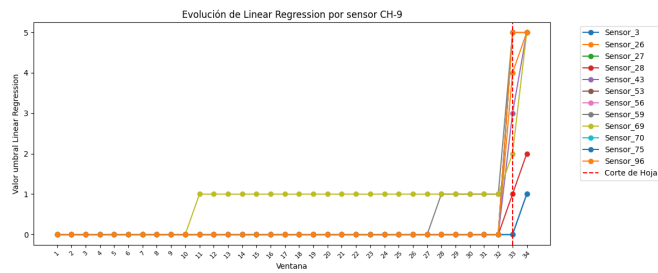


XGB CH-8

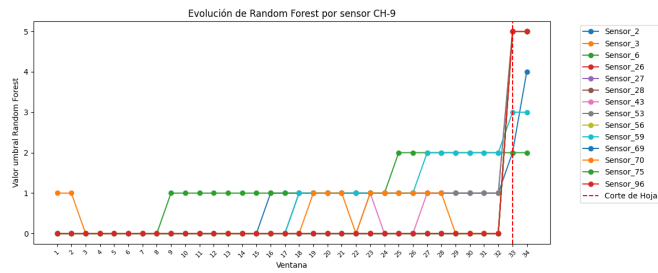
Figura A1.8: Comparación vertical de métodos en CH-8



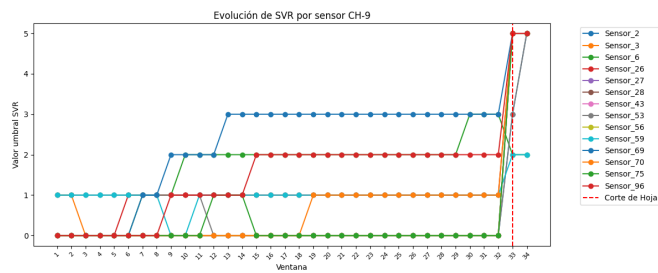
KNN CH-9



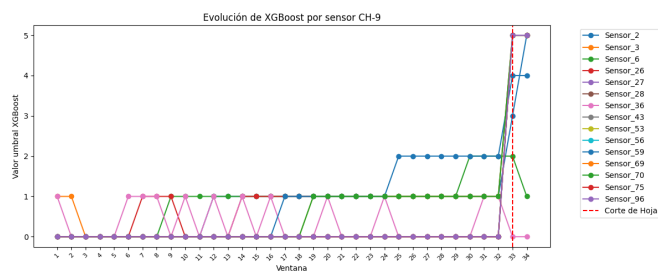
LR CH-9



RF CH-9

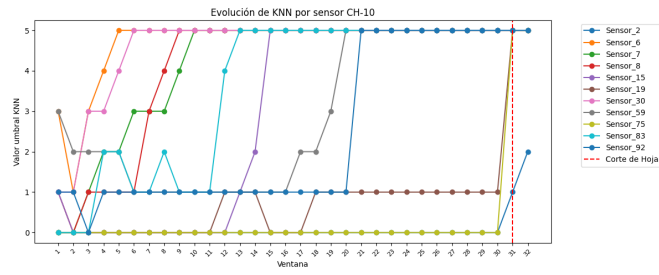


SVR CH-9

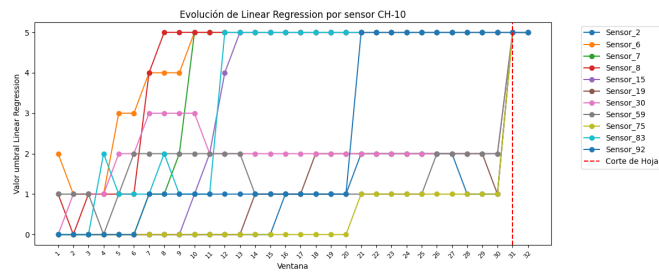


XGB CH-9

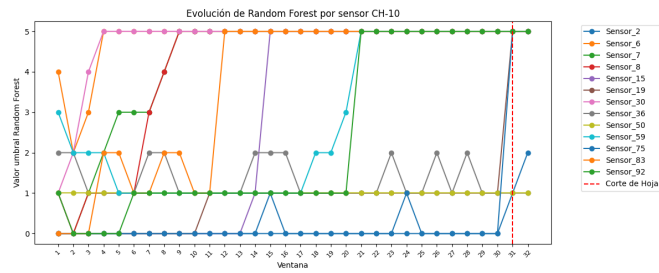
Figura A1.9: Comparación vertical de métodos en CH-9



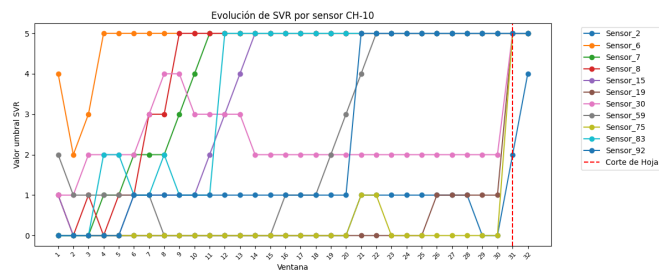
KNN CH-10



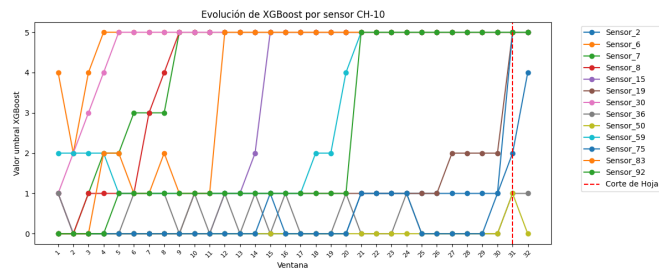
LR CH-10



RF CH-10

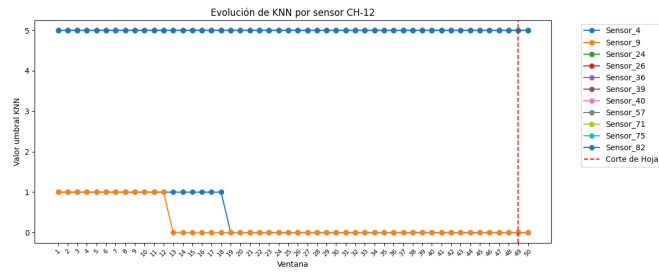


SVR CH-10

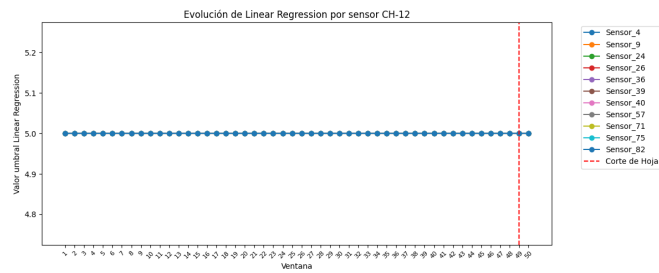


XGB CH-10

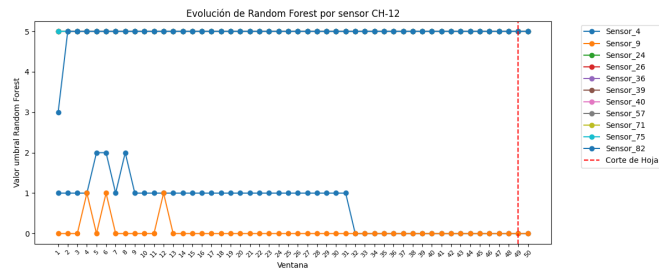
Figura A1.10: Comparación vertical de métodos en CH-10



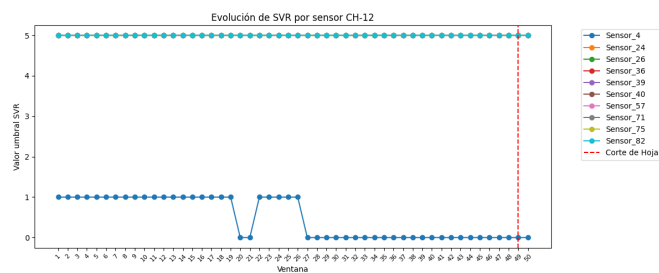
KNN CH-12



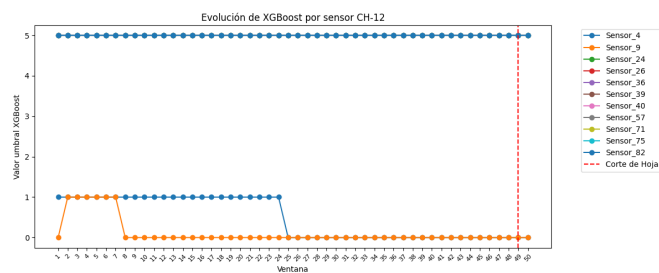
LR CH-12



RF CH-12

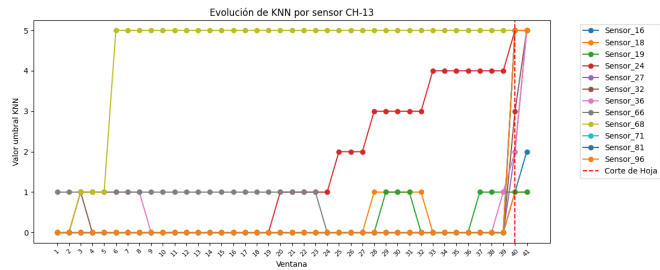


SVR CH-12

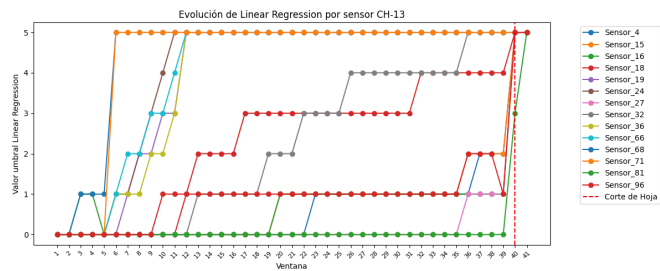


XGB CH-12

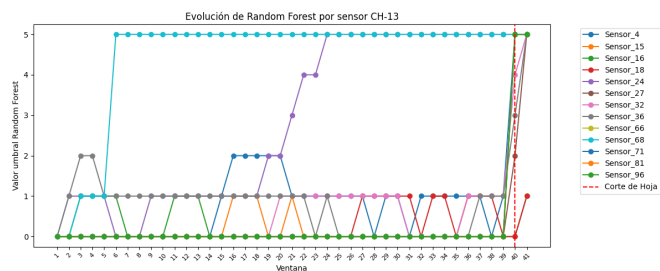
Figura A1.12: Comparación vertical de métodos en CH-12



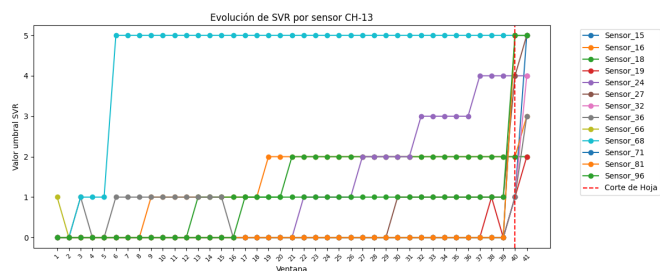
KNN CH-13



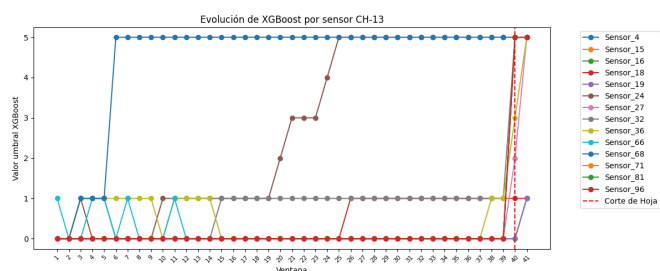
LR CH-13



RF CH-13

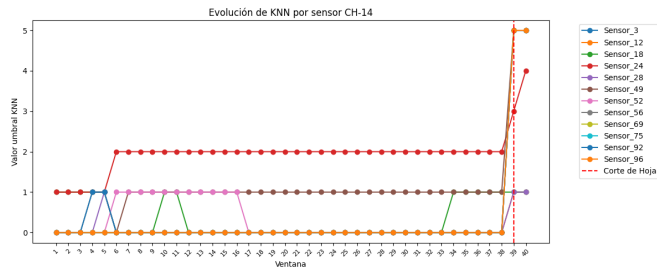


SVR CH-13

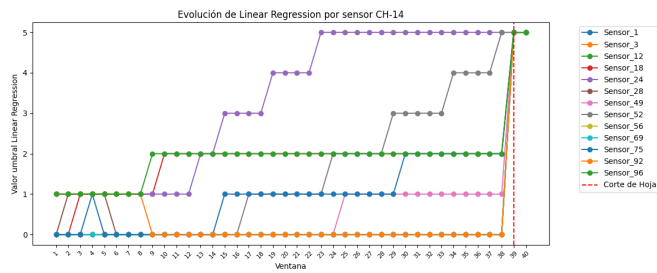


XGB CH-13

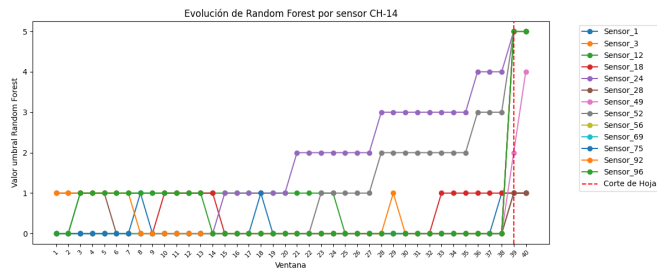
Figura A1.13: Comparación vertical de métodos en CH-13



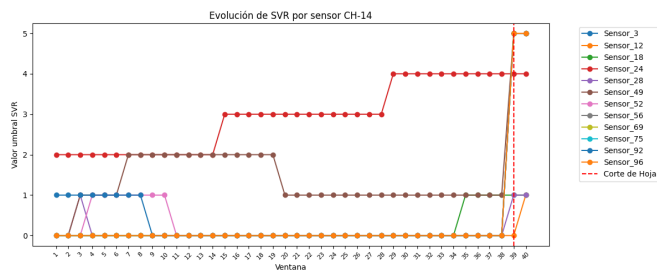
KNN CH-14



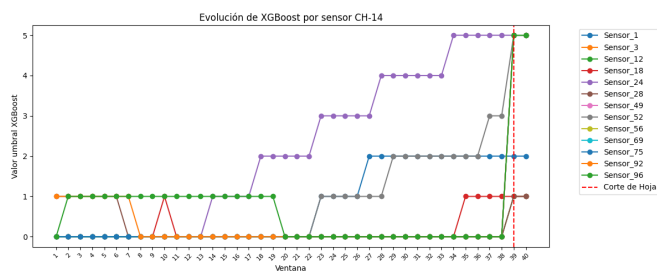
LR CH-14



RF CH-14



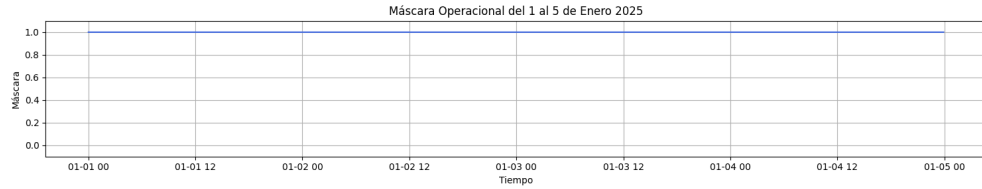
SVR CH-14



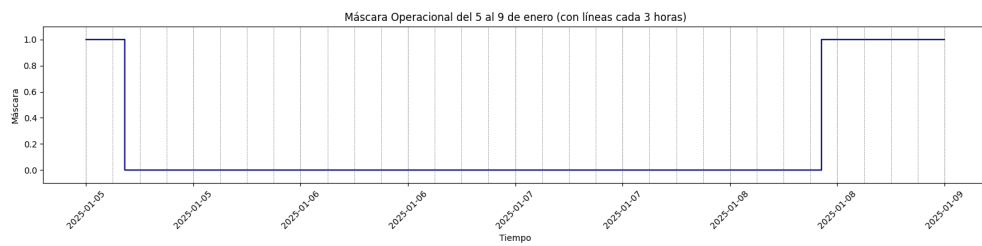
XGB CH-14

Figura A1.14: Comparación vertical de métodos en CH-14

A2. Máscaras booleanas de entrenamiento y test



Máscara booleana train CH-1

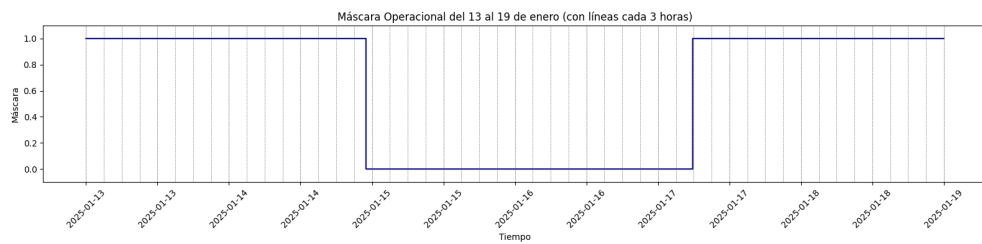


Máscara booleana test CH-1

Figura A2.1: Máscaras booleanas CH-1



Máscara booleana train CH-2

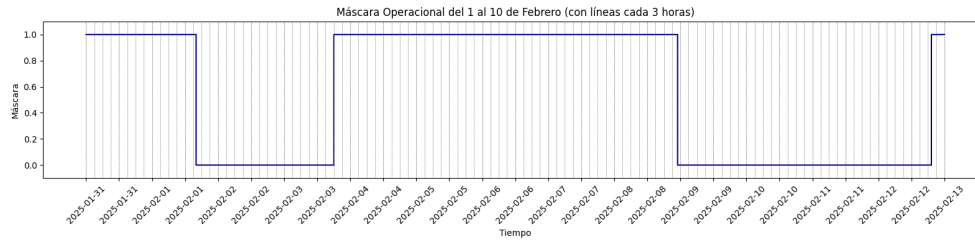


Máscara booleana test CH-2

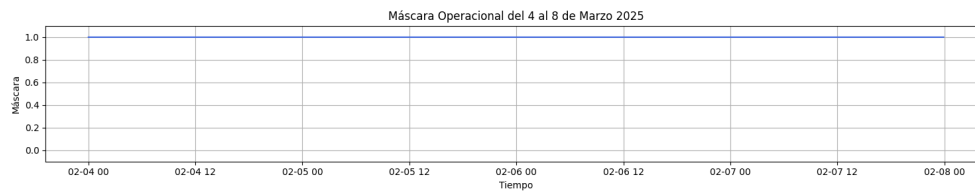
Figura A2.2: Máscaras booleanas CH-2



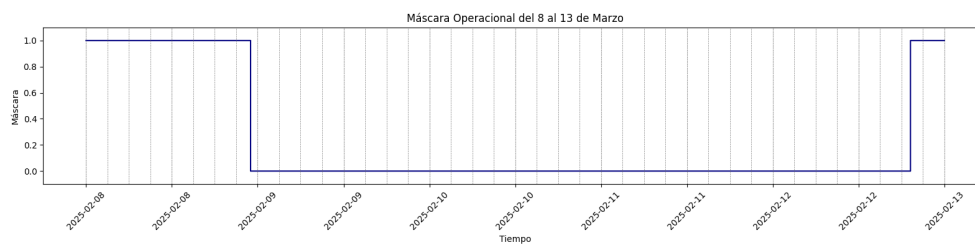
Máscara booleana train CH-3



Máscara booleana test CH-3

Figura A2.3: Máscaras booleanas CH-3

Máscara booleana train CH-4

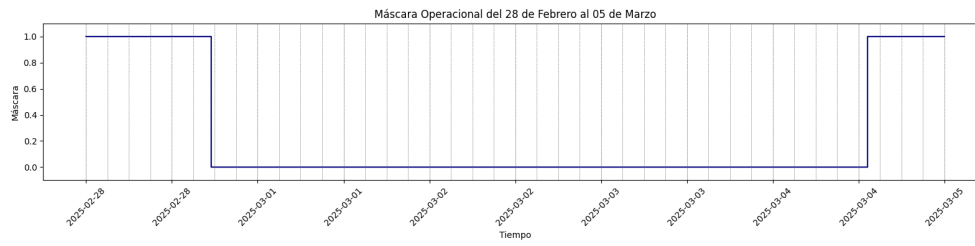


Máscara booleana test CH-4

Figura A2.4: Máscaras booleanas CH-4



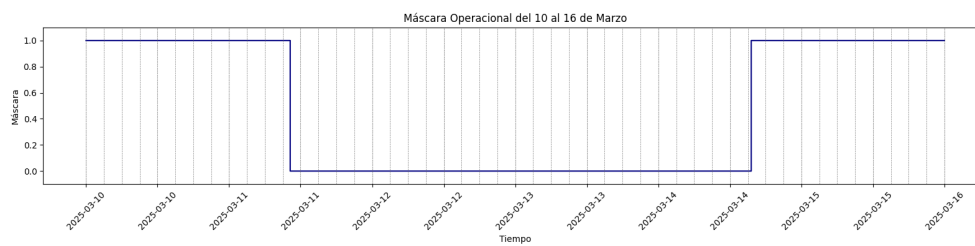
Máscara booleana train CH-7



Máscara booleana test CH-7

Figura A2.7: Máscaras booleanas CH-7

Máscara booleana train CH-8

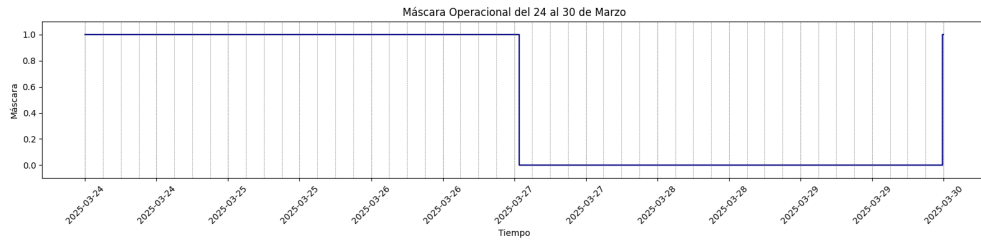


Máscara booleana test CH-8

Figura A2.8: Máscaras booleanas CH-8



Máscara booleana train CH-9

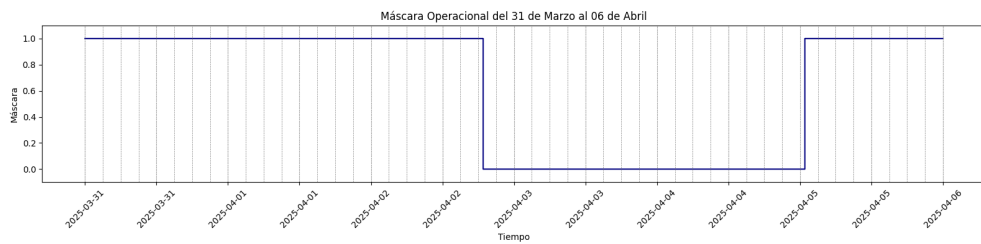


Máscara booleana test CH-9

Figura A2.9: Máscaras booleanas CH-9



Máscara booleana train CH-10

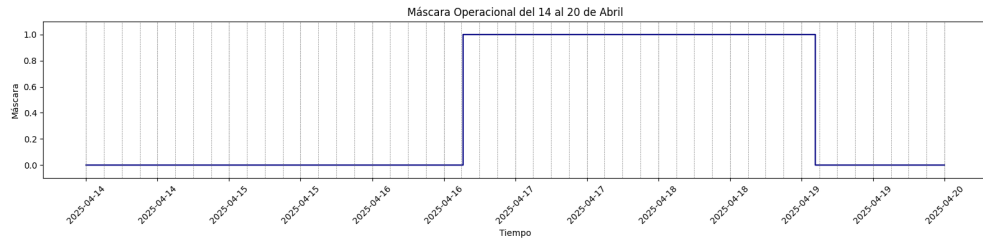


Máscara booleana test CH-10

Figura A2.10: Máscaras booleanas CH-10



Máscara booleana train CH-11

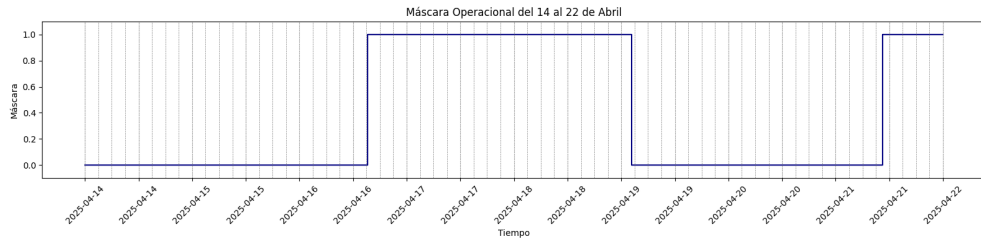


Máscara booleana test CH-11

Figura A2.11: Máscaras booleanas CH-11



Máscara booleana train CH-12

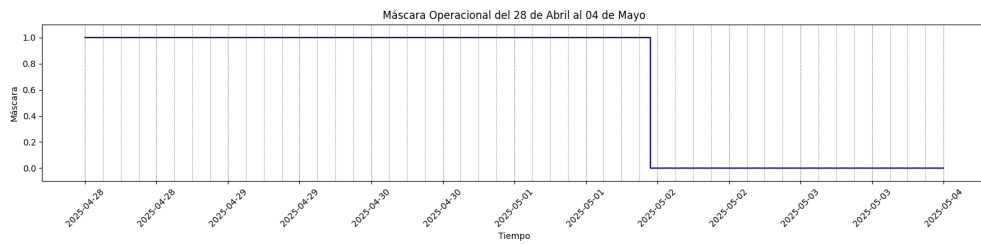


Máscara booleana test CH-12

Figura A2.12: Máscaras booleanas CH-12



Máscara booleana train CH-13

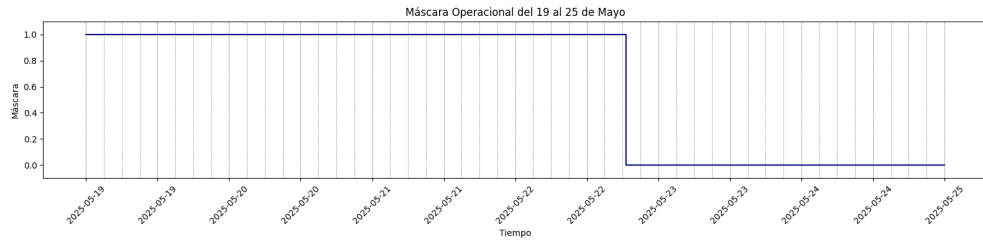


Máscara booleana test CH-13

Figura A2.13: Máscaras booleanas CH-13



Máscara booleana train CH-14



Máscara booleana test CH-14

Figura A2.14: Máscaras booleanas CH-14