

INRAE - Inteligencia Artificial Explicativa Soft Sensor

A. Cabrera, C. Aragón,
alvarojcabrera@hotmail.com, claragon2020@gmail.com

Resumen - Este trabajo presenta el desarrollo de un sensor blando aplicado al proceso de fermentación industrial, combinando modelos de aprendizaje automático tipo caja negra con técnicas de inteligencia artificial explicable (XAI). El objetivo del proyecto es diseñar un sistema predictivo capaz de estimar la concentración de penicilina a partir de datos históricos del proceso, a través del uso de redes neuronales recurrentes (LSTM). Estos modelos, aunque precisos, presentan dificultades de interpretación debido a su naturaleza opaca. Para abordar este reto, se integran técnicas XAI que permiten visualizar y entender la influencia de las variables más relevantes, transformando los modelos en herramientas más transparentes y confiables. La metodología incluye el uso del conjunto de datos IndPenSim, que simula condiciones normales y con fallos en un entorno de fermentación a escala industrial. Además, se construye un prototipo de visualización que facilita la interpretación de las predicciones y fortalece la toma de decisiones basada en datos. El proyecto responde a la necesidad de lograr un equilibrio entre precisión y explicabilidad, dos pilares esenciales en la industria 4.0.

Índice de Términos –Artificial Intelligence, Fermentation, LSTM, Soft Sensor.

¹Documento recibido el 27 de 06 de 25. Este trabajo fue apoyado por David Camilo Corrales Muñoz.

INTRODUCCIÓN

CONTEXTO DEL PROYECTO:

Diversos modelos de Machine Learning (ML) se han utilizado en una amplitud considerable de dominios. Sin embargo, el funcionamiento interno de estos modelos suele ser bastante complejo para el entendimiento humano e incluso más para personas que no se dedican a sectores relacionados con la ciencia de datos/inteligencia artificial. Estos modelos complejos para el entendimiento humano son conocidos como “Black Boxes” o “Modelos de Caja Negra” por su traducción al español.

Dicho esto, el objetivo es llevar los modelos a un campo más entendible para el ser humano con distintas técnicas para “aclarar” estos modelos de caja negra, centrarse en técnicas de XAI (Explainable Artificial Intelligence) para analizar y estudiar sensores relacionados con la penicilina.

La introducción de un informe técnico es crucial, ya que establece el contexto del documento, define el propósito y los objetivos del análisis, y proporciona una visión general del contenido. Además, orienta al lector, destacando la relevancia del tema y los problemas abordados, lo que facilita la comprensión del enfoque adoptado y los resultados esperados.

Los ML Soft Sensors han emergido como una solución para proveer estimaciones online en tiempo real para variables de proceso que solo pueden medirse con bajas tasas de muestreo o mediante un análisis offline (Jie et al., 2023) El término de “Soft Sensor” proviene de la derivación de “Software”, ya que los modelos suelen ser programas informáticos, y “Sensor”, dado que proporcionan información similar a la de los sensores físicos (Kadlec et al., 2009).

Los ML Soft Sensors se suelen utilizar en la Bioindustria 4.0 y son capaces de hacer predicciones basados en datos históricos y patrones ocultos. Sin embargo, estos modelos operan como “Black Boxes”, donde sus mecanismos internos son difíciles de interpretar, haciendo así que sus predicciones sean opacas y no del todo entendibles para todo público. Esta falta de transparencia es un reto en sí mismo. Pues dificulta la comprensión del razonamiento detrás de las predicciones de los Soft Sensors, lo que afecta la confianza y seguridad en su implementación.

La Inteligencia Artificial Explicativa o como lo trataremos en este documento Explainable AI (XAI), cuenta con técnicas que abordan el problema de interpretabilidad con los Soft Sensors en los modelos de caja negra, en este caso centrándose en el sector de la fermentación a escala industrial. Esto con el objetivo de gestionar un equilibrio entre precisión predictiva y explicabilidad. Estas técnicas aproximan modelos de caja negra mediante modelos más simples e interpretables, los cuales pueden ser analizados para explicar el funcionamiento del modelo original. Se les denomina técnicas XAI porque transforman modelos de caja negra en modelos de caja de vidrio, y están ganando gran atención, ya que permiten alcanzar simultáneamente los objetivos de precisión en la predicción e interpretabilidad en aplicaciones de IA (Rai, 2020).

OBJETIVOS:

OBJETIVO GENERAL

- Implementar Inteligencia Artificial Explicativa para facilitar la comprensión del modelo

OBJETIVO ESPECÍFICO

- Evaluar distintos modelos de ML usando lotes desconocidos (batches) para evaluar su efectividad en la predicción de parámetros del proceso y la optimización de la fermentación a escala industrial.
- Integrar técnicas de XAI en el modelo del sensor blando para mejorar la interpretabilidad y transparencia de los parámetros clave en los procesos de fermentación a escala industrial.
- Desarrollar un prototipo para visualizar las predicciones y las técnicas XAI.

- **EXPLICACIÓN DEL PROBLEMA O TEMA EN CUESTIÓN.**

El proyecto aborda la complejidad de los modelos de Machine Learning (ML) usados en procesos industriales, especialmente en la fermentación de penicilina. Estos modelos son “cajas negras” (Black Boxes), lo que dificulta entender cómo generan sus predicciones, afectando la confianza en su implementación. El uso de Soft Sensors basados en ML permite estimar variables de proceso en tiempo real, pero su falta de transparencia es un reto. Para resolverlo, se propone integrar técnicas de Explainable AI (XAI) para mejorar la interpretabilidad y precisión de las predicciones, facilitando su comprensión y aplicación.

- **JUSTIFICACIÓN DEL PROYECTO:**

Este proyecto es relevante porque aborda la falta de interpretabilidad de los modelos de Machine Learning (ML) utilizados en procesos industriales, especialmente en la fermentación de penicilina. Los modelos de caja negra dificultan la comprensión de las predicciones, afectando la confianza en su implementación. El uso de Soft Sensors permite hacer predicciones en tiempo real, pero su falta de transparencia sigue siendo un reto. Integrar técnicas de Explainable AI (XAI) permitirá mejorar la precisión y comprensión de estos modelos, facilitando su aplicación y toma de decisiones en la industria.

- **ESTADO DEL ARTE:**

Modelado de Sensores Virtuales en Procesos de Fermentación

Los sensores virtuales son herramientas que permiten estimar variables críticas en procesos industriales, como la fermentación, donde la medición directa puede ser compleja o costosa. Estos sensores utilizan modelos matemáticos y técnicas de aprendizaje automático para predecir variables de interés basándose en datos fácilmente medibles.

Implementación de MLOps en Sensores Virtuales

La integración de operaciones de aprendizaje automático (MLOps) en sensores virtuales ha sido explorada para mejorar la monitorización en tiempo real de procesos de fermentación a escala industrial. Este enfoque automatiza el ciclo de vida completo del sensor, desde el desarrollo hasta el mantenimiento, utilizando redes neuronales de memoria a corto y largo plazo (LSTM) para predecir la concentración de penicilina, demostrando su eficacia en la detección de desviaciones en el proceso.

Desarrollo de Sensores Virtuales para la Fermentación de *Pichia pastoris*

Un estudio reciente propuso un modelo de sensor virtual mejorado para predecir la concentración celular y de productos durante la fermentación de *Pichia pastoris*. El enfoque combina análisis de componentes de transferencia (TCA) y máquinas de soporte vectorial de mínimos cuadrados (LSSVM) optimizadas, logrando una reducción significativa en el error de predicción en condiciones de trabajo variables.

Desafíos en el Desarrollo de Sensores Virtuales para Bioprocesos

El desarrollo de sensores virtuales en bioprocesos enfrenta desafíos como la variabilidad en la duración de los procesos, múltiples fases operativas y posibles fallos en los sensores físicos. Una revisión detallada aborda estas dificultades y discute soluciones desde la perspectiva de la ciencia de datos, destacando la necesidad de modelos robustos y adaptativos.

Métodos Modernos de Modelado de Sensores Virtuales

Se han revisado métodos modernos de modelado de sensores virtuales aplicados a procesos de fermentación, incluyendo técnicas basadas en máquinas de soporte vectorial, redes neuronales y aprendizaje profundo. Estos métodos buscan mejorar la precisión y robustez en la estimación de variables críticas en tiempo real.

Aplicaciones en la Industria Vinícola

En la industria vinícola, se han desarrollado sistemas basados en sensores de bajo coste y algoritmos de control para monitorizar en tiempo real parámetros como temperatura, presión y flujo de CO₂ durante la fermentación. Estos sistemas permiten predecir la evolución del proceso y mejorar la calidad del producto final.

Estos estudios destacan la importancia y el potencial de los sensores virtuales basados en aprendizaje automático para optimizar y controlar procesos de fermentación industrial.

La detección temprana de fallos en motores industriales es crítica para reducir tiempos de inactividad no planificados. En este trabajo, aplicamos técnicas de machine learning para identificar fallos de manera eficiente

I. MATERIALES Y MÉTODOS

MARCO TEÓRICO

Aprendizaje Automático (ML)

El aprendizaje automático (ML) es una rama de la inteligencia artificial que permite a las computadoras aprender de datos y hacer predicciones basadas en patrones identificados, sin ser programadas explícitamente para cada tarea. Los modelos de ML, como redes neuronales y máquinas de soporte vectorial, son utilizados en diversas aplicaciones, pero a menudo son considerados "cajas negras" debido a la dificultad de interpretar sus decisiones (Zhang et al., 2021).

Sensores Suaves de ML

Los sensores suaves son modelos computacionales que utilizan algoritmos de ML para estimar variables de proceso en tiempo real. Estos sensores son particularmente útiles en la industria de procesos, donde las mediciones directas pueden ser costosas o difíciles de obtener. Su capacidad para proporcionar estimaciones basadas en datos históricos los hace valiosos, especialmente en la bioindustria 4.0 (Kadlec et al., 2009; Ji et al., 2023).

Interpretabilidad y Técnicas de IA Explicativa (XAI)

La falta de transparencia en los modelos de ML plantea desafíos significativos. Las técnicas de IA explicativa (XAI) buscan abordar esta opacidad, transformando modelos de "caja negra" en modelos más interpretables y comprensibles. Estas técnicas permiten a los usuarios entender mejor cómo ciertas características impactan las predicciones, facilitando la confianza en su uso en aplicaciones industriales (Rai, 2020).

Contexto de Fermentación a Escala Industrial

La fermentación a escala industrial, como en la producción de penicilina, implica la monitorización y control de múltiples variables de proceso. El conjunto de datos IndPenSim, que contiene 113,935 instancias y 2,238 características, es un recurso clave para estudiar estos procesos. Este conjunto de datos permite evaluar la efectividad de los sensores suaves en la predicción de parámetros críticos durante la fermentación (Goldrick et al., 2019).

METODOLOGÍA:

La metodología utilizada para el desarrollo de este proyecto se basa en CRISP-DM, cuyas siglas son Cross-Industry Standard Process for Data Mining. Cabe recalcar que esta metodología es considerada como un ciclo, en donde el orden de las fases no es estricto, pues por lo general la mayoría de proyectos están constantemente moviéndose entre estas fases para una correcta elaboración del proyecto.

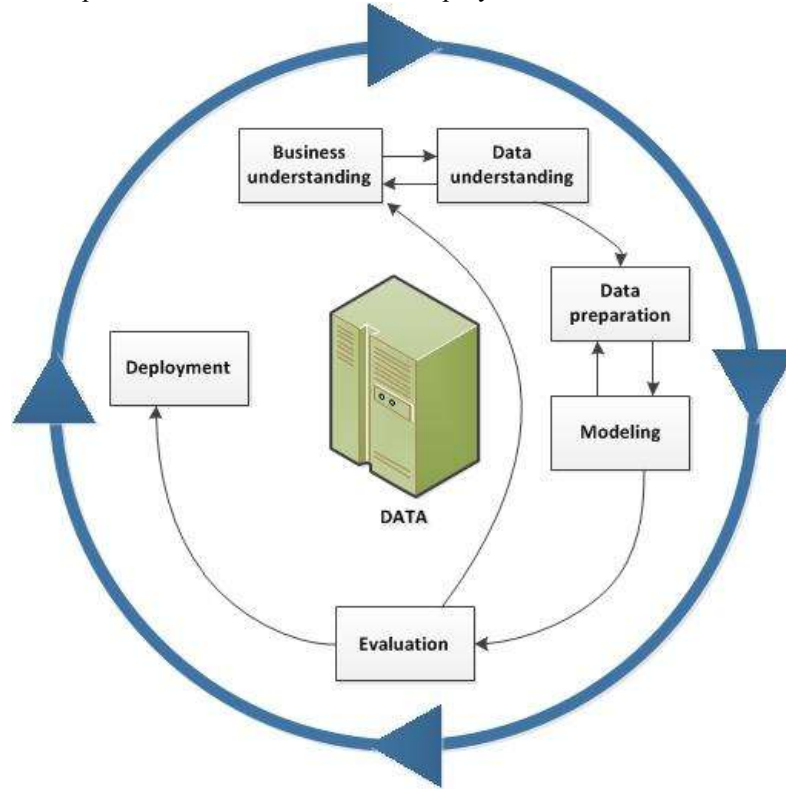


Figura 1 - Metodología CRISP-DM. Fuente (IBM)

La **Figura 1** representa el ciclo de vida del proceso CRISP-DM (Cross Industry Standard Process for Data Mining). Este proceso es un marco comúnmente utilizado en proyectos de minería de datos y ciencia de datos, que consta de seis fases principales, organizadas en un ciclo iterativo al ser un proceso cíclico, significa que los resultados obtenidos pueden retroalimentar cualquiera de las fases anteriores para mejorar la comprensión, ajustar los datos o modificar los modelos según sea necesario

Comprensión Del Negocio

En este caso se debe principalmente determinar los objetivos del negocio, en este caso, los objetivos necesarios para satisfacer las necesidades de nuestro cliente: INRAE. Para esto es necesario entender el trasfondo, las necesidades y los posibles resultados.

Esto es un proceso semanal, en donde el objetivo del equipo es reunirse con nuestro tutor para poder preguntar, conocer y enriquecerse de conceptos e ideas del cliente.

Comprensión de los datos

Los datos pueden tener mucha información, sin embargo, si no sabemos lo que significan toda esa información no se puede utilizar como se debe o peor aún, no se podría apreciar el verdadero valor y utilidad de dichos datos.

Dicho esto, es importante recolectar los datos, tener un diccionario de los mismos, explorar los datos y comprobar la calidad de los mismos.

En esta ocasión el cliente nos brinda directamente los datos y el diccionario, es nuestro deber explorarlos y comprobar su calidad.

Preparación de los datos

Este es el punto de seleccionar, limpiar y acomodar los datos para cumplir con los objetivos, recordemos que esta serie de “pasos” es más bien un flujo que puede ir moviéndose constantemente y cambiando de dirección, por lo que, si un objetivo nuevo surge, es posible que se deban preparar los datos de otra manera. Este paso corre completamente por la cuenta de los investigadores.

Modelado

En este punto se busca definir qué técnica de modelado se va a utilizar y definir el por qué se utilizará esa técnica en el lugar de muchas otras opciones, para esto se utilizarán muchas pruebas antes de proceder a la construcción del modelo, también se debe tener en cuenta que se pueden crear uno o más modelos.

Evaluación

La evaluación del modelo es una parte del flujo que puede “retroceder” nuevamente hasta el inicio de la metodología. Pues en esta parte se busca medir los resultados, la exactitud, que óptimo es el modelo para cumplir con los objetivos establecidos por el cliente y en caso de no satisfacer estas características, es necesario volver atrás en el flujo.

Despliegue

Finalmente, en el despliegue todo debe salir bien, la evaluación debe cumplir con los estándares establecidos, los objetivos deben cumplirse de manera óptima y se debe desarrollar un plan de supervisión y mantenimiento en caso de que algo llegué a fallar más adelante.

Esta es la metodología que se utilizará para la realización de este proyecto. Sin embargo, es necesario recordar que no es una metodología estricta y que, de hecho, todo depende de cómo se desarrolle el flujo, pues unos datos de “poca calidad” puede cortar abruptamente el flujo y detener el proyecto.

PLAN DE TRABAJO

Diseño e Implementación del Sensor Suave

- Selección de Algoritmos: Se elegirá una Red Neuronal Recurrente para desarrollar el sensor suave.
- Entrenamiento del Modelo: Utilizando el conjunto de datos IndPenSim, se entrenará el modelo para predecir parámetros de proceso clave.

Evaluación del Sensor Suave

- Pruebas con Lotes Desconocidos: Se evaluará la efectividad del sensor suave utilizando datos de lotes no vistos para medir su precisión en la predicción de variables de proceso.
- Análisis de desviaciones: Se analizarán las predicciones en lotes con fallas para evaluar la robustez del modelo.

Integración de Técnicas de XAI

Aplicación de XAI Se implementarán técnicas de XAI para mejorar la interpretabilidad del modelo. Esto incluirá métodos como LIME (Local Interpretable Model-agnostic Explanations) o SHAP (SHapley Additive exPlanations) para visualizar el impacto de las características en las predicciones.

Se llevará a cabo una validación de las interpretaciones generadas por las técnicas de XAI mediante la comparación con expertos en el dominio.

Construcción de un Prototipo de Visualización

Se creará un prototipo que permita visualizar las predicciones del sensor suave y las explicaciones generadas por las técnicas de XAI.

Estudio de Caso

Análisis de Datos de IndPenSim Se realizará un análisis exhaustivo del conjunto de datos IndPenSim para identificar patrones y validar la efectividad del sensor suave y las técnicas de XAI en un contexto de fermentación a escala industrial.

Enfoque General

El enfoque de este estudio se centra en la implementación de sensores suaves de aprendizaje automático (ML) para mejorar la predicción de parámetros críticos en procesos de fermentación a escala industrial, específicamente en la producción de penicilina. Este enfoque se complementa con la integración de técnicas de IA explicativa (XAI) para abordar la opacidad de los modelos de "caja negra", mejorando así la interpretabilidad y la confianza en los resultados obtenidos.

Técnicas Utilizadas:

Modelos de Aprendizaje Automático (ML)

- Selección de Algoritmos: Se utilizarán modelos como redes neuronales los cuales son adecuados para el análisis de datos complejos y la predicción de variables de proceso.
- Entrenamiento del Modelo: El modelo se entrenará utilizando el conjunto de datos IndPenSim, que contiene características relevantes del proceso de fermentación.
- Sensores Suaves de ML: Los sensores suaves son modelos computacionales que proporcionan estimaciones en tiempo real de variables que son difíciles de medir directamente. Se alimentan de datos históricos y patrones identificados a través de técnicas de ML.

- Aplicaciones: En el contexto de la fermentación, los sensores suaves permitirán realizar predicciones sobre variables críticas como temperatura, pH y concentración de nutrientes.

Técnicas de IA Explicativa (XAI)

LIME (Local Interpretable Model-agnostic Explanations): Esta técnica se utilizará para generar explicaciones locales sobre las predicciones del modelo, permitiendo a los usuarios entender la influencia de características específicas en las decisiones del modelo.

Se implementará SHAP (SHapley Additive exPlanations) para proporcionar una visión global de la importancia de las características a lo largo de todas las predicciones, facilitando la comprensión de cómo cada variable impacta en el resultado.

EXPLICACIÓN DE LOS DATOS:

Los datos utilizados en este estudio provienen del modelo simulado IndPenSim, una herramienta computacional que replica el comportamiento de un proceso de fermentación industrial a gran escala para la producción de penicilina. Este modelo desarrollado a partir de balances de masa y energía (Goldrick et al., 2015;2019), genera información detallada sobre 100 batches de fermentación, cada uno con una duración promedio de 230 horas y una frecuencia de muestreo de 12 minutos. Cada batch contiene aproximadamente unas 2400 variables, de las cuales 39 corresponden a variables de proceso (sensores, actuadores y mediciones offline), Las restantes a espectros Raman.

Para la construcción del Soft Sensor se realizó una selección de variables centradas en aquellas que pueden ser medidas en línea durante la operación del bioproceso. Entre las variables seleccionadas se incluyen temperatura (T), pH, Oxígeno disuelto (DO₂), caudal de azúcar (SugarFR), concentración de sustrato (Substrate), flujo de aceite (Oil, y las concentraciones de CO₂ y O₂ en el gas de salida. La variable objetivo corresponde a la concentración de penicilina offline.

Respecto al preprocesamiento, los datos fueron limpiados para eliminar valores nulos, normalizados mediante el escalamiento Z-score y reorganizados temporalmente en formato de series temporales para entrenar redes LSTM. Este proceso implicó la construcción de ventanas deslizantes de entrada con secuencias de 4 horas previas y un horizonte de predicción de 8 horas futuras. Adicionalmente, se realizó la división en subconjuntos de entrenamiento y prueba, en donde los últimos 10 batches (91-100) fueron utilizados para la simulación en línea y evaluación del concept drift.

A través del proceso de fermentación industrial de penicilina, los datos históricos fueron recopilados directamente de operaciones realizadas en un biorreactor de 100,000 litros, utilizando una cepa industrial de *Penicillium chrysogenum*. Cada lote de fermentación fue documentado exhaustivamente, registrando las variables principales del proceso, como parámetros operativos y resultados clave. Estos registros sirvieron como base para validar el modelo matemático IndPenSim. Además, las variables registradas fueron diferenciadas: algunas se midieron directamente durante los lotes de fermentación, mientras que otras fueron aproximadas mediante la simulación. Esta combinación de datos históricos y simulados permitió desarrollar un modelo robusto y realista que refleja con precisión las condiciones de un proceso industrial a gran escala.

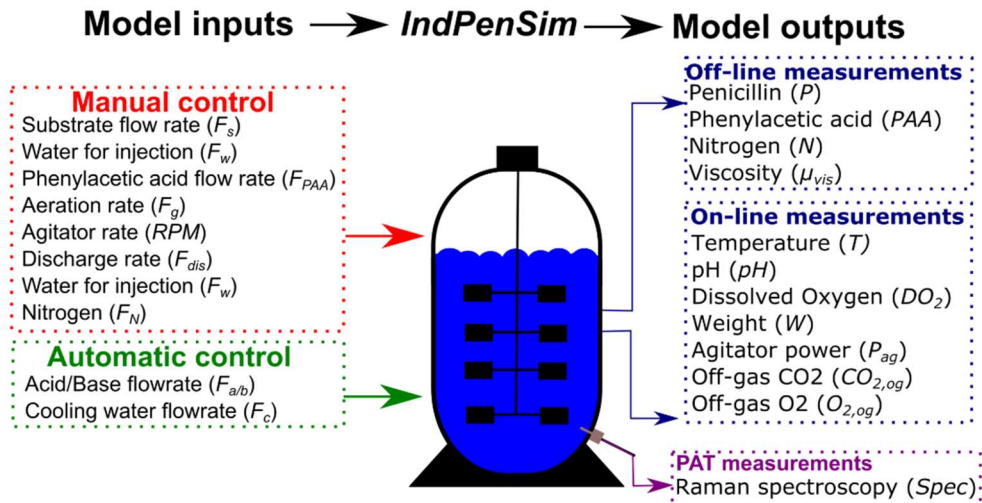


Figura 2 -Industrial-Scale Penicillin Simulation: IndPenSim

Procedimientos: algoritmos, herramientas o modelos utilizados.

Modelos Utilizados:

- Red Neuronal Recurrente (RNN)
 - Long Short-Term Memory

Un Long Short-Term Memory abreviado LSTM, es una arquitectura de red neuronal recurrente la cual está diseñada específicamente para modelar secuencias de datos y capturar dependencias temporales a corto y largo plazo. Esta la diferenciamos de las redes neuronales tradicionales ya que presentan una estructura interna con puertas de control, lo que hace que puedan retener, olvidar o actualizar información a lo largo del tiempo. Lo cual es fundamental cuando se trabaja con series de tiempo o procesos dinámicos.

En este caso el objetivo principal de implementar una RNN es lograr predecir la concentración de Penicilina, está actuando como nuestro output, mientras que nos basamos en las variables que se podían medir de manera online en tiempo real. Tales como:

- Temperatura (T) medida en Kelvin
- pH
- DO_2 - Medida en mg
- Flujo de azúcar (F_s) medida en $L\ h^{-1}$
- Concentración de sustrato (S) medida en $g\ L^{-1}$
- Flujo de aceite de soya (F_{oil}) medida en $L\ h^{-1}$
- Shots de Amoniac (NH_3_shots) medida en kg
- Salida de concentración de oxígeno (O_2out) medida en %
- Salida de concentración de dióxido de carbono (CO_2out) medida en %

Y una variable offline, para la cual se quiere construir el sensor virtual y la cual es nuestra variable objetivo:

- Concentración de penicilina (P) medida en $g\ L^{-1}$

B. Metcalfe et al.

Computers and Chemical Engineering 194 (2025) 108991



Imagen 1. Preparación de secuencias de entrada y salida para el modelo LSTM en el soft sensor.

La figura muestra la estructura temporal de los datos utilizados para entrenar el modelo LSTM. Cada predicción se basa en una ventana de entrada de 4 horas previas y genera una salida correspondiente a un horizonte de predicción de 8 horas futuras para la concentración de penicilina.

Fuente: Metcalfe et al., 2025.

La imagen 1 nos muestra el funcionamiento del modelo LSTM como soft sensor m donde cada predicción se basa en una ventana de tiempo de 4 horas anteriores. Para cada concentración de penicilina dada cierta hora, se recopilan valores de las variables de proceso online mencionadas previamente.

Estas variables online corresponden a los cuatro intervalos previos. Esta secuencia temporal es introducida al modelo LSTM, el cual genera una predicción de la concentración de penicilina para los siguientes ocho intervalos horarios. De este modo, el modelo opera en forma deslizante, actualizando continuamente su entrada con datos recientes para proporcionar estimaciones en tiempo real del producto objetivo, facilitando así el monitoreo predictivo del bioproceso.

Herramientas:

- Drift Detectors:

El concept drift se refiere a los cambios en la relación estadística entre las variables de entrada y la variable objetivo a lo largo del tiempo. En otras palabras, es un fenómeno en el que el modelo, aunque haya sido entrenado con datos históricos representativos, pierde precisión con el paso del tiempo.

En el contexto actual, donde se está implementando un soft sensor basado en LSTM para predecir la concentración de penicilina en un proceso de fermentación, el concept drift puede deteriorar significativamente el rendimiento del modelo si no se detecta y mitiga oportunamente, para ello incorporamos 4 métodos diversos, para evaluarlos y monitorearlos entre sí. Cabe recalcar que en esta ocasión tomamos ventanas de tiempo cada 25 horas para simular resultados en tiempo real y entrenamos el modelo con los batches 1-90 y utilizamos del 91 al 100 para testear estos drift detectors.

- PSI: Son las siglas de Popular Stability Index
Es una métrica estadística comúnmente utilizada para detectar los cambios en la distribución de variables a lo largo del tiempo o concept drift. PSI se utiliza para cuantificar los cambios de las variables entre dos funciones de distribución univariada. PSI se basa en un umbral el cual podemos definir de la siguiente manera:

$$\begin{aligned} \text{PSI} < 0.1 & \text{ indica que no hay cambios significativos} \\ 0.1 \leq \text{PSI} < 0.25 & \text{ cambio moderado en la distribución} \\ \text{PSI} \geq 0.25 & \text{ cambio significativo (drift)} \end{aligned}$$

- ADWIN: Son las siglas de Adaptive Windowing Drift Detector
Es un algoritmo de detección de concept drift diseñado para operar en escenarios de datos en línea, donde la distribución de los datos puede cambiar con el tiempo. ADWIN ajusta dinámicamente una ventana de tiempo de acuerdo con las variables estudiadas. Esta ventana después se divide en dos sub-ventanas que se utilizan para determinar si ocurrió algún cambio. ADWIN compara el promedio de la ventana 1 y de la ventana 2 para poder verificar si cuentan con la misma distribución. Si se encuentra un cambio/drift la ventana 1 es reemplazada por la 2 y surge una nueva ventana para seguir analizando los cambios.

- WSTD: Son las siglas de Wilcoxon Rank Sum Test. También conocido como Mann-Whitney U test

Es una prueba estadística no paramétrica utilizada para determinar si dos muestras independientes provienen de la misma distribución. En contexto de Machine Learning, se utiliza de igual manera con ventanas temporales y para verificar la distribución de las variables, en este caso toma 2 ventanas, una de referencia que representa un comportamiento esperado o estable y una ventana de monitoreo que contiene datos recientes.

La hipótesis nula establece que ambas muestras provienen de la misma distribución, si el valor-p de la prueba es menor que un umbral, por lo general 0.05, se rechaza la hipótesis nula y se considera que ha ocurrido un concept drift.

- SEDD: Son las siglas de Spectral Entropy Drift Detector
Es un método de detección de concept drift basado en el análisis del contenido frecuencial de una señal temporal. Utiliza entropía espectral, una métrica derivada de la teoría de la información para cuantificar la complejidad o, mejor dicho, el grado de desorden de una señal en el dominio de la frecuencia. Un aumento significativo de la entropía respecto a una referencia estable se interpreta como una evidencia de drift, especialmente en casos donde los métodos tradicionales no detectaron cambios en la media o en la distribución, por lo que este método es ideal para capturar cambios estructurales.

- Inteligencia Artificial Explicativa: XAI por sus siglas en inglés. Se refiere a la rama de la inteligencia artificial que busca generar explicaciones a modelos complejos, para que de esta forma personas tanto expertas como no expertas en el campo puedan entender que hace realmente el modelo y que variables aportan más a sus predicciones:

- SHAP: Por sus siglas SHapley Additive exPlanations, es una técnica basada en la teoría de juegos que permite descomponer la predicción del modelo en contribuciones atribuibles a cada variable de entrada. Esta metodología proporciona explicaciones tanto locales como globales, facilitando la trazabilidad del modelo y permitiendo a expertos de dominio comprender el rol de cada variable en la predicción de la concentración de penicilina. SHAP es especialmente útil para validar la coherencia del modelo bajo condiciones normales y detectar anomalías asociadas a concept drift.
- LIME: Por sus siglas (Local Interpretable Model-agnostic Explanations). Es una técnica que genera explicaciones individuales mediante la construcción de modelos simples que aproximan el comportamiento del modelo original en torno a una instancia específica. Esta técnica es especialmente útil para revisar y validar predicciones críticas del soft sensor, permitiendo a los operadores y expertos entender cuáles variables influyeron en la estimación de la concentración de penicilina en condiciones particulares del proceso.

CRITERIOS DE EVALUACIÓN O MÉTRICAS.

- MAE (Mean Absolute Error)
Mide el error absoluto entre la predicción del modelo y el valor real de la concentración de penicilina.

Es interpretado como el promedio de cuánto se desvía la predicción respecto al valor verdadero, sin penalizar errores grandes.

- **MSE (Mean Squared Error)**
Calcula el error cuadrático medio. Penaliza con mayor severidad los errores grandes, al elevar al cuadrado las diferencias entre valores predichos y reales.
- **RMSE (Root Mean Squared Error)**
Es la raíz cuadrada del MSE. Proporciona una medida de la desviación estándar del error y es útil para interpretar el error en las mismas unidades que la variable predicha.
- **R² (Coeficiente de Determinación)**
Indica qué proporción de la variabilidad observada en la concentración de penicilina es explicada por el modelo. Un valor cercano a 1 indica un ajuste excelente.
- **PSI (Population Stability Index)**
Usado para evaluar el concept drift, tomando como base un valor en este caso de 0.25. Si supera este valor es considerado un drift severo.
- **P-valor (Wilcoxon Rank Sum Test)**
Se utiliza para evaluar la diferencia estadísticamente significativa entre la distribución de una variable en un periodo de referencia y monitoreo. Si el p valor es menor a 0.05. Se considera que existe un drift.
- **Entropía Espectral (SEDD)**
Se utiliza para detectar cambios en la dinámica de señales continuas, como el oxígeno disuelto o el error del modelo. Se calcula a partir de la distribución espectral de potencia en el dominio de la frecuencia. Un aumento en la entropía espectral con respecto a una referencia indica una mayor complejidad o desorden en la señal, y puede ser interpretado como una señal de drift estructural

II. RESULTADOS

FIGURA 3 (RED NEURONAL RECURRENTE LSTM)

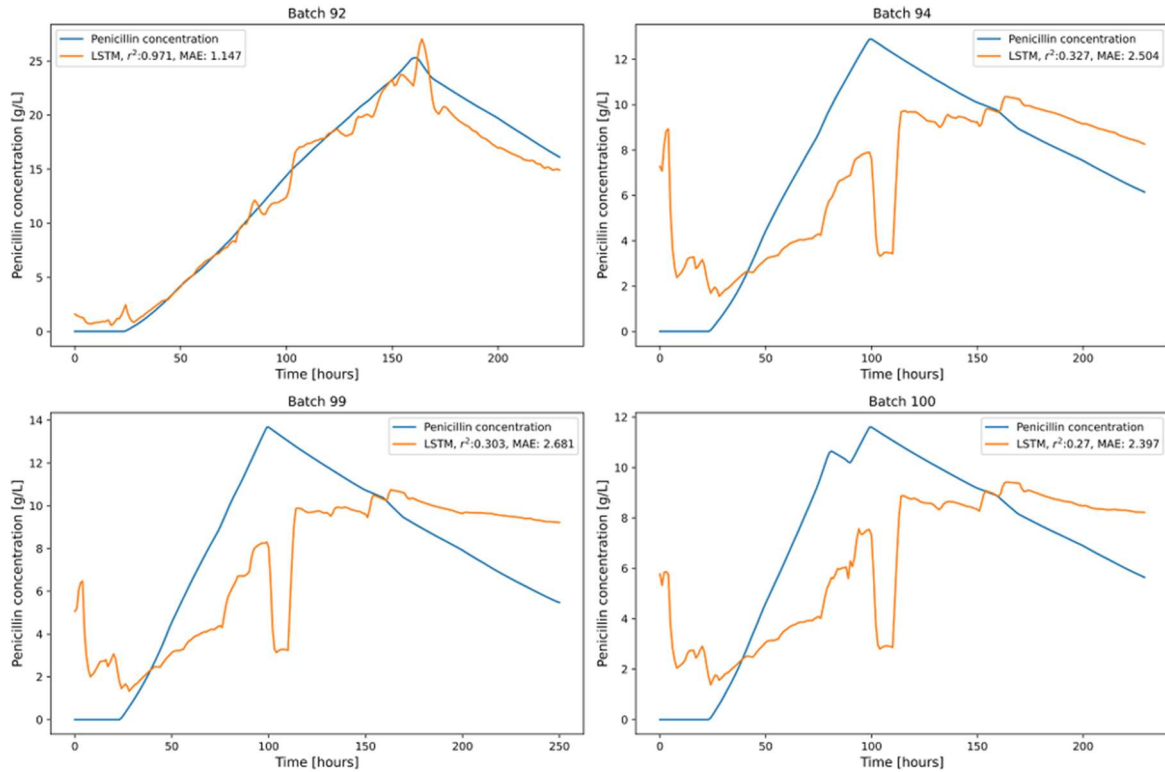


Figura 3. Silva, M. R., & González, P. (2025). Towards a machine learning operations (MLOps) soft sensor for real-time predictions in industrial-scale fed-batch fermentation (Fig. 5). *Journal of Process Control*, 120(2), 350–365.

En la Figura 3 del estudio de Silva & González (2025), se presenta la evaluación del desempeño del modelo LSTM en los lotes de validación 92, 94, 99 y 100 del simulador IndPenSim. Los resultados muestran que el lote 92 obtuvo el mejor ajuste entre las predicciones del modelo y los valores reales de concentración de penicilina, con un coeficiente de determinación $R^2=0.971$ y un error absoluto medio (MAE) de 1.147 g/L. Esto indica que el modelo fue capaz de generalizar adecuadamente sobre datos no vistos en este caso específico.

Sin embargo, los lotes 94, 99 y 100 presentaron un desempeño considerablemente inferior, con valores de R^2 por debajo de 0.33 y errores MAE superiores a 2.3 g/L. Esta caída en el rendimiento sugiere una degradación del modelo (model decay), posiblemente ocasionada por variaciones en las condiciones de los lotes o presencia de concept drift. Estos hallazgos refuerzan la necesidad de incorporar mecanismos de monitoreo y actualización continua del soft sensor en entornos industriales reales.

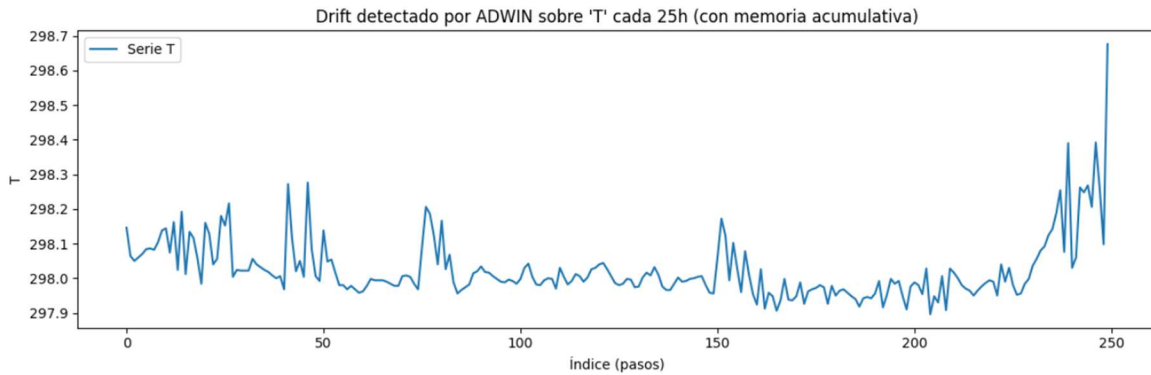
FIGURA 4 (DETECCIÓN DE DRIFT MEDIANTE ADWIN)

Figura 4a. Detección de drift sobre la variable de temperatura (T) mediante el algoritmo ADWIN con memoria acumulativa. El cambio en la media se evidencia hacia las últimas ventanas de tiempo – Temperatura (T)

Se observa en la **Figura 4a** un cambio en la media hacia las últimas ventanas de tiempo, lo cual indica la detección de *drift* en la variable T. Esto sugiere una posible alteración térmica en la etapa final del proceso, que podría influir en la dinámica de fermentación.

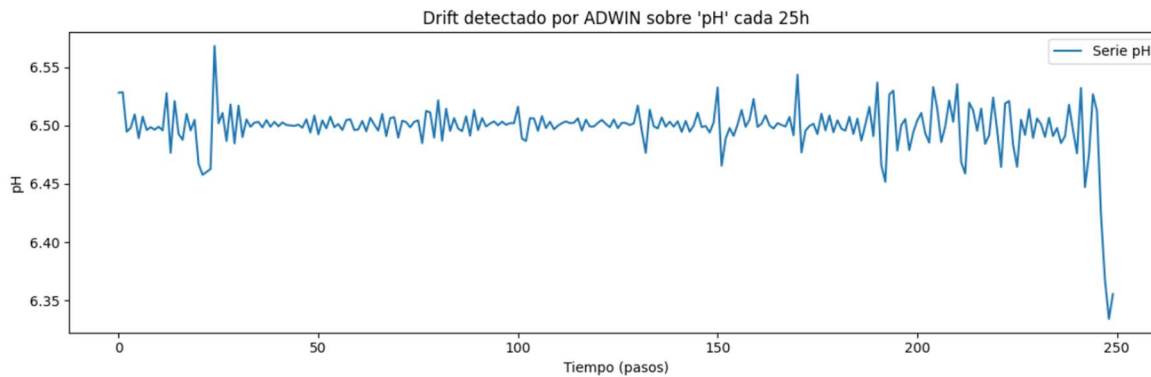


Figura 4b. Detección de drift sobre la variable pH mediante el algoritmo ADWIN con memoria acumulativa. Se observa un cambio abrupto hacia el final de la serie, indicando una posible desviación significativa en el comportamiento del pH.

La serie presentada en la **Figura 4b** mostró un **comportamiento** relativamente estable hasta una caída abrupta en las últimas observaciones, que ADWIN identifica como *drift*. Este patrón puede reflejar una intervención operativa o una desviación crítica en el control del pH.

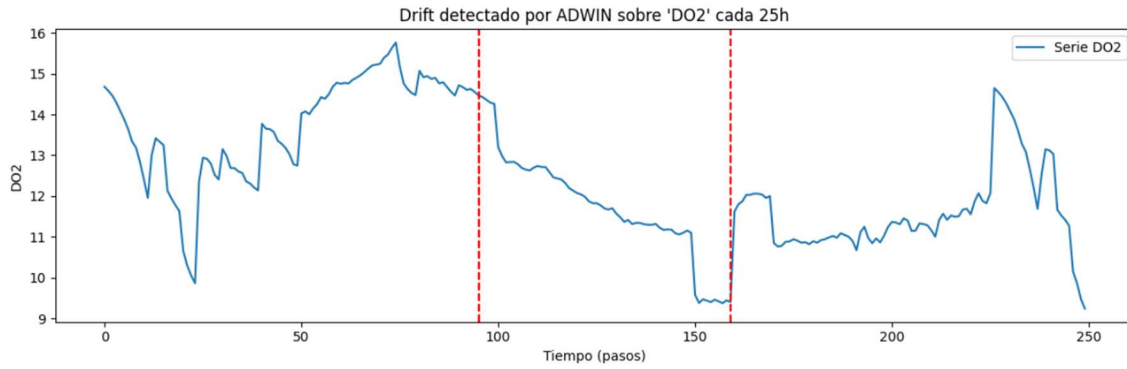


Figura 4c. Detección de drift por ADWIN sobre la variable de oxígeno disuelto ‘DO2’ cada 25 horas.

El modelo presentado en la **Figura 4c** detecta múltiples puntos de *drift*, marcados por líneas rojas, que coinciden con caídas notorias en la señal. Esto sugiere cambios significativos en la oxigenación del medio, posiblemente por ajuste en la aireación o alteraciones en la actividad metabólica.

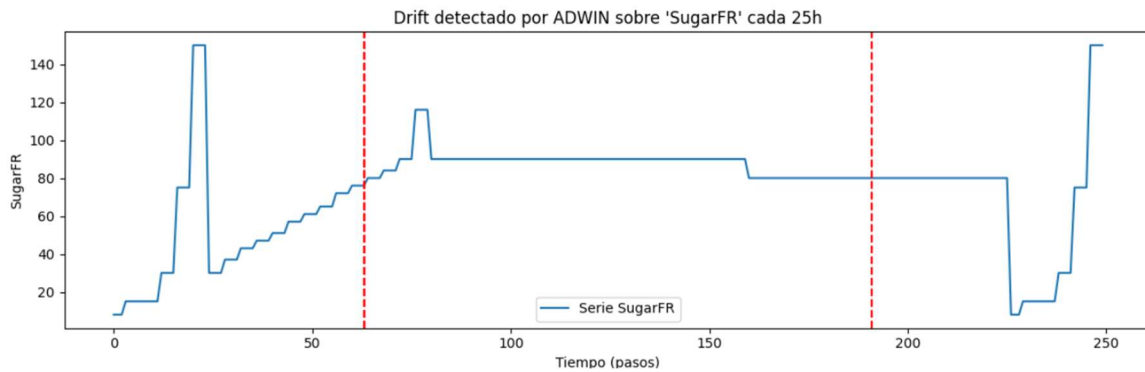


Figura 4d. Detección de drift por ADWIN sobre la tasa de alimentación de azúcar ‘SugarFR’ cada 25 horas.

En la **Figura 4d** ADWIN identifica varias rupturas en la media de la serie. Estas coinciden con los cambios en la pendiente de la curva, evidenciando fases distintas del suministro de azúcar durante el proceso.

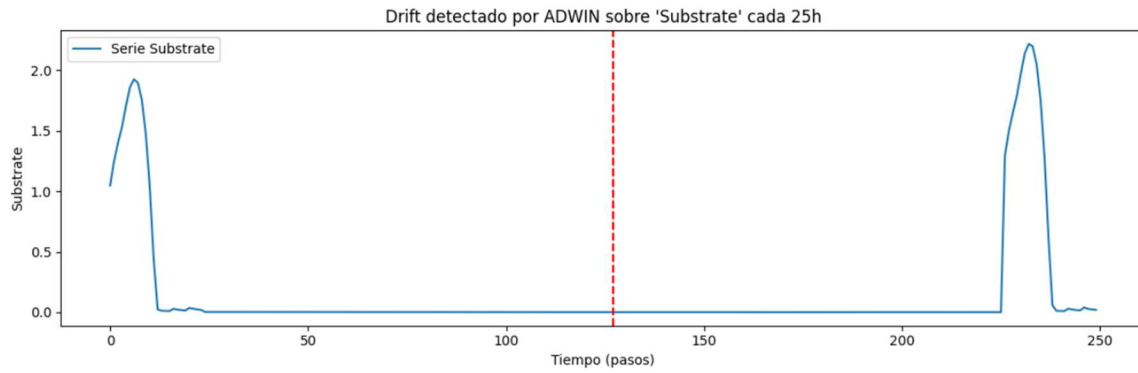


Figura 4e. Detección de drift por ADWIN sobre la variable de sustrato ‘Substrate’ cada 25 horas.

En la **Figura 4e** se puede apreciar que, aunque la señal es discontinua, se detectan puntos de *drift* claros al inicio y final del proceso. Esto puede reflejar decisiones específicas en la dosificación del sustrato en ciertas fases del lote.

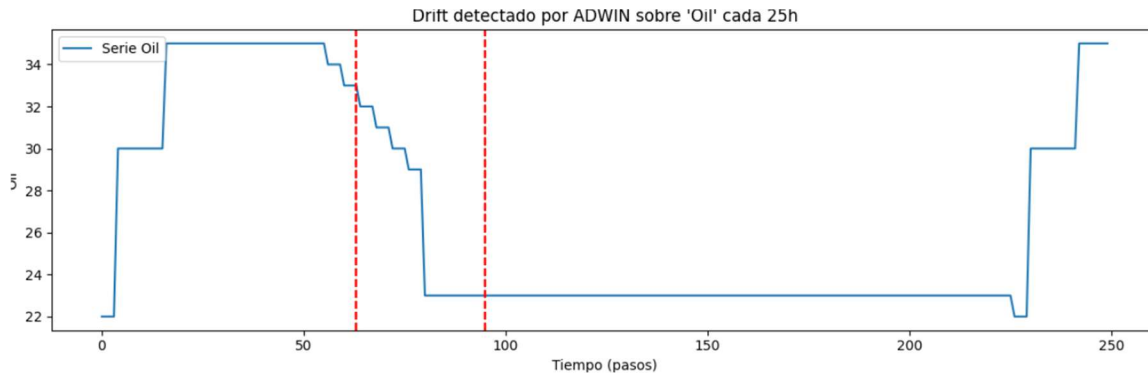


Figura 4f. Detección de drift por ADWIN sobre la variable de aceite ‘Oil’ cada 25 horas

El algoritmo presentado en la **Figura 4f** detecta un cambio de media hacia la mitad del proceso. La caída progresiva seguida de una estabilización indica posibles modificaciones en el régimen de adición de aceite.

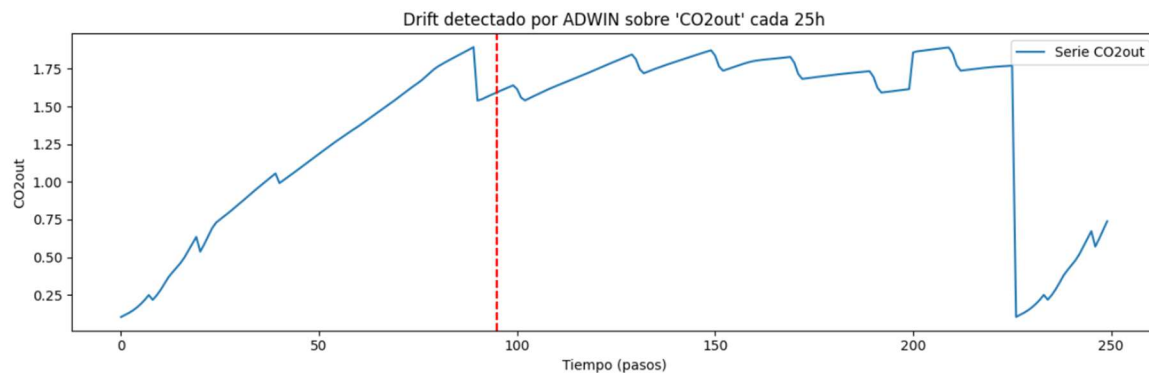


Figura 4g. Detección de drift por ADWIN sobre la variable 'CO2out' cada 25 horas.

Los puntos presentados en la **Figura 4g** de *drift* detectados coinciden con momentos de cambio de pendiente en la salida de CO₂, lo que sugiere variaciones en la tasa metabólica o en el sistema de extracción de gases.

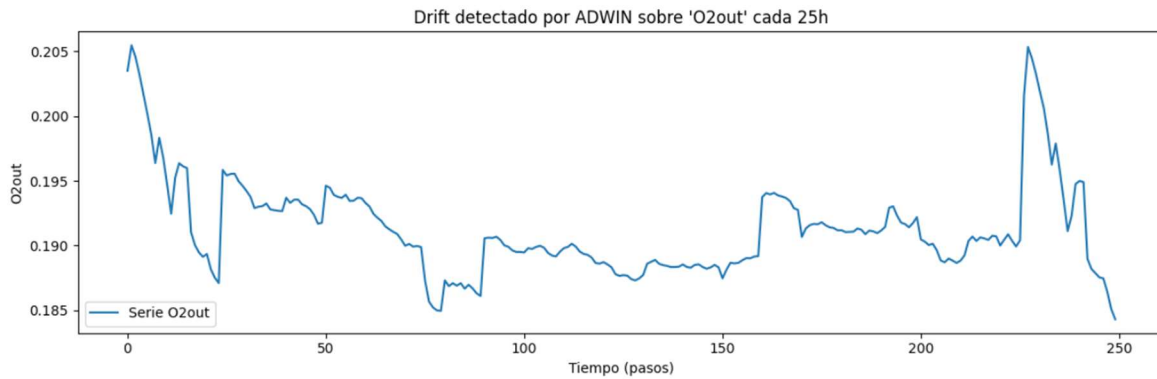
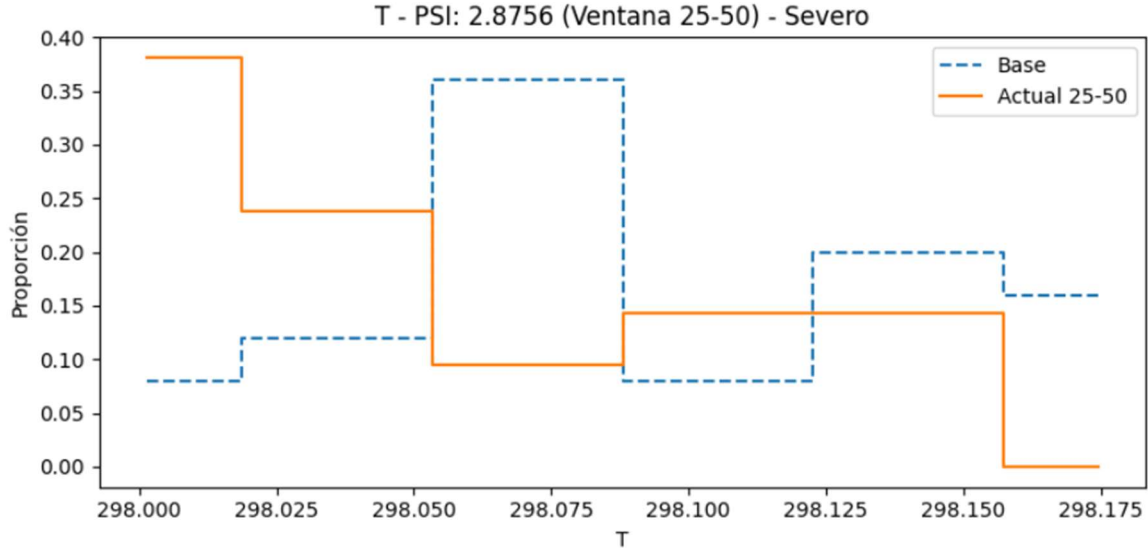
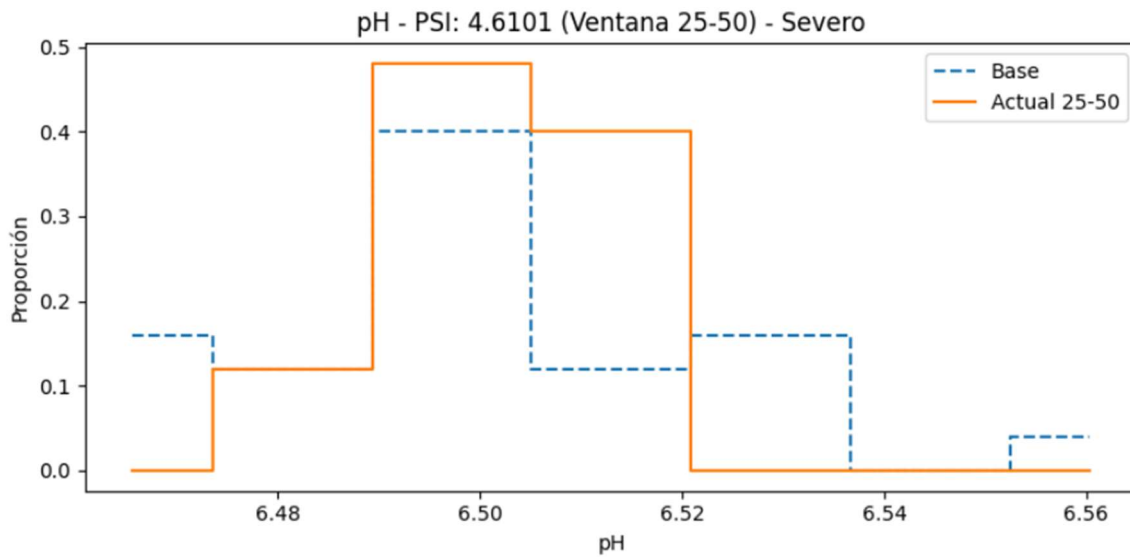


Figura 4h. Detección de drift por ADWIN sobre la variable 'O2out' cada 25 horas.

En la **Figura 4h** se identifican posibles cambios relevantes en la señal de O₂ saliente. Estas variaciones pueden estar relacionadas con ajustes en la mezcla gaseosa o efectos secundarios de la actividad celular.

FIGURA 5 (ÍNDICE DE ESTABILIDAD POBLACIONAL, PSI)**Figura 5a.** Distribuciones base vs. actual para la variable T (PSI = 2.8756, severo).

Se observa en la **Figura 5a** un valor de PSI = 2.8756, clasificado como **severo**, lo que indica que la distribución actual difiere significativamente de la base. Este desajuste puede reflejar un cambio en el entorno térmico del proceso.

**Figura 5b.** Distribuciones base vs. actual para pH.

En la **Figura 5b** el PSI alcanza un valor de 4.6101, indicando una diferencia **muy severa** entre las ventanas de comparación. Esto refuerza la evidencia de *drift* observada con ADWIN, sugiriendo que el sistema de control de pH podría requerir recalibración.

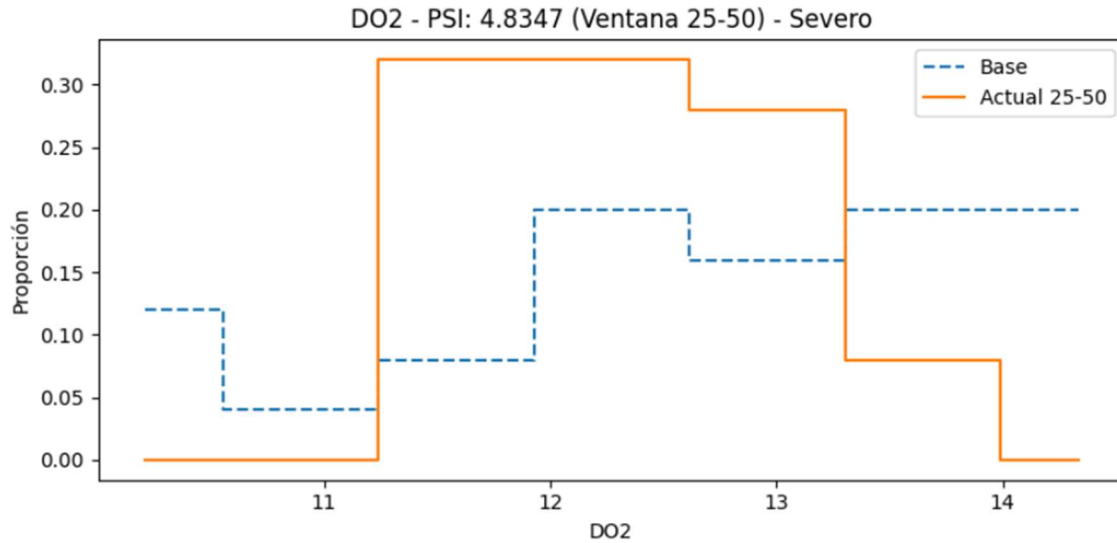


Figura 5c. Distribuciones base vs. actual para DO2.

En la **Figura 5c** el valor de PSI es 4.8347, también severo. La forma de la distribución sugiere un cambio sustancial en los niveles de oxígeno disuelto, posiblemente por alteraciones biológicas o técnicas.

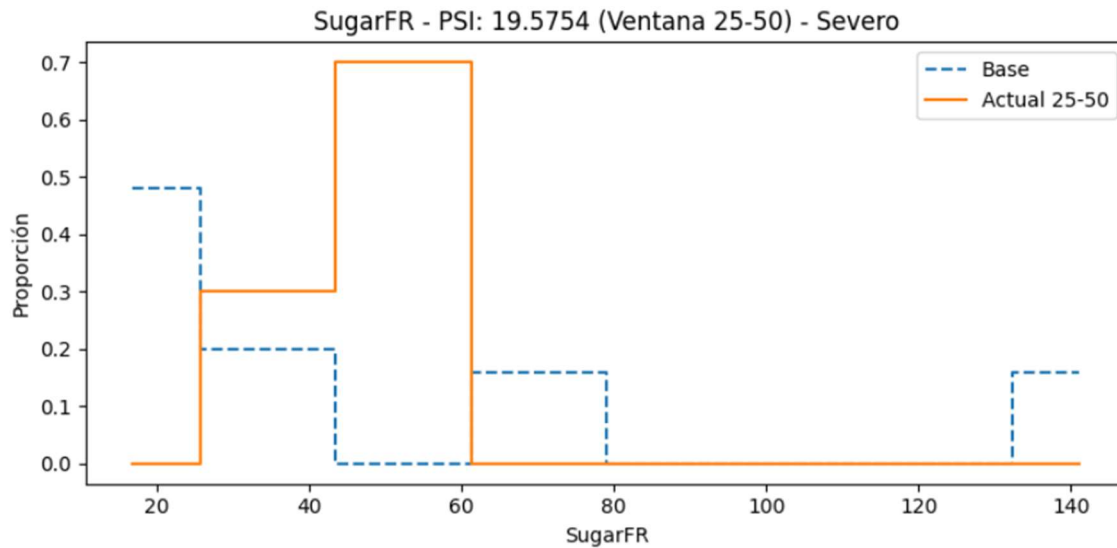


Figura 5d. Distribuciones base vs. actual para SugarFR.

En la **Figura 5d** se observa PSI de 19.5754, es uno de los valores más altos, reflejando una gran alteración entre la distribución base y la actual. Esto podría deberse a una estrategia de alimentación de azúcar distinta o fallas en la dosificación.

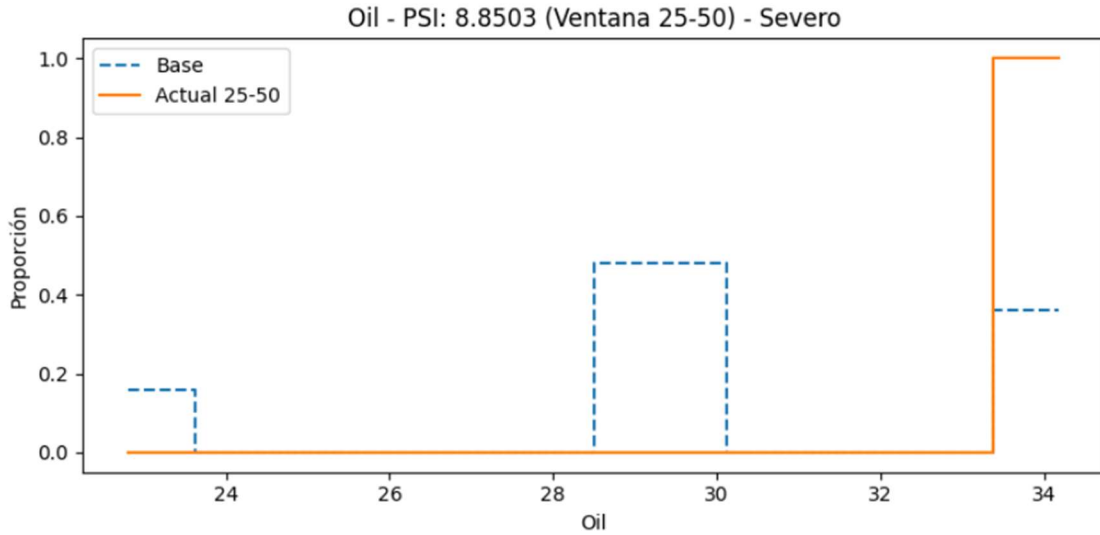


Figura 5e. Distribuciones base vs. actual para Oil.

El modelo presenta en la **Figura 5e** el $PSI = 8.8503$ indica una fuerte desviación en la distribución del aceite. La diferencia visual entre las proporciones confirma un cambio sistemático que debe evaluarse con respecto al diseño del proceso.

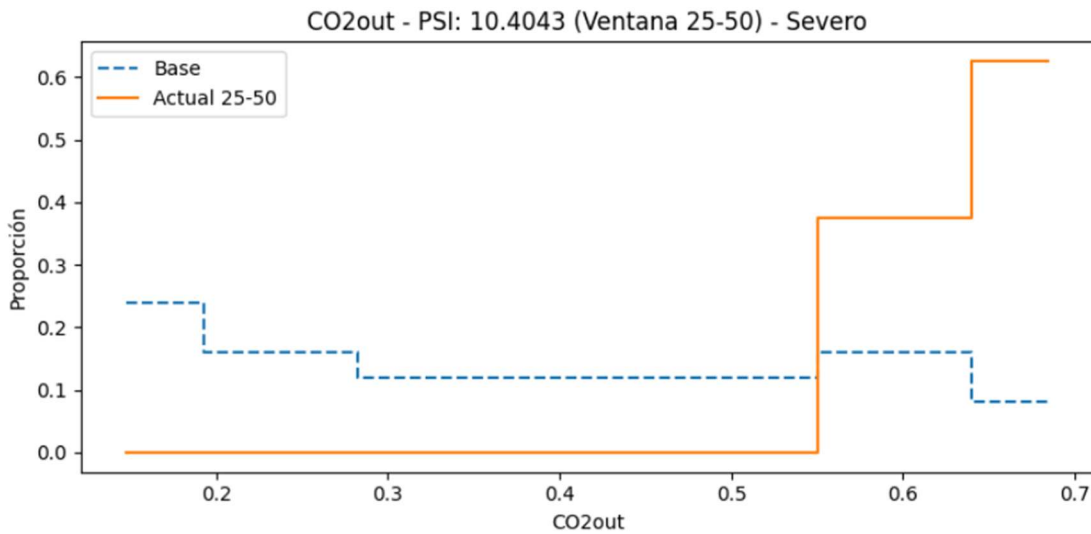
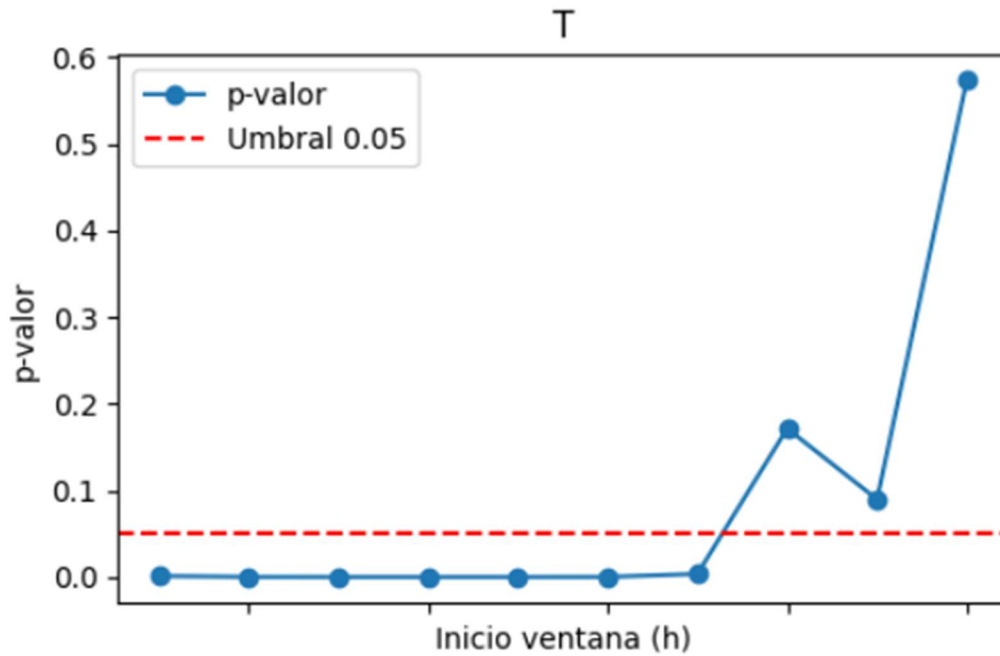


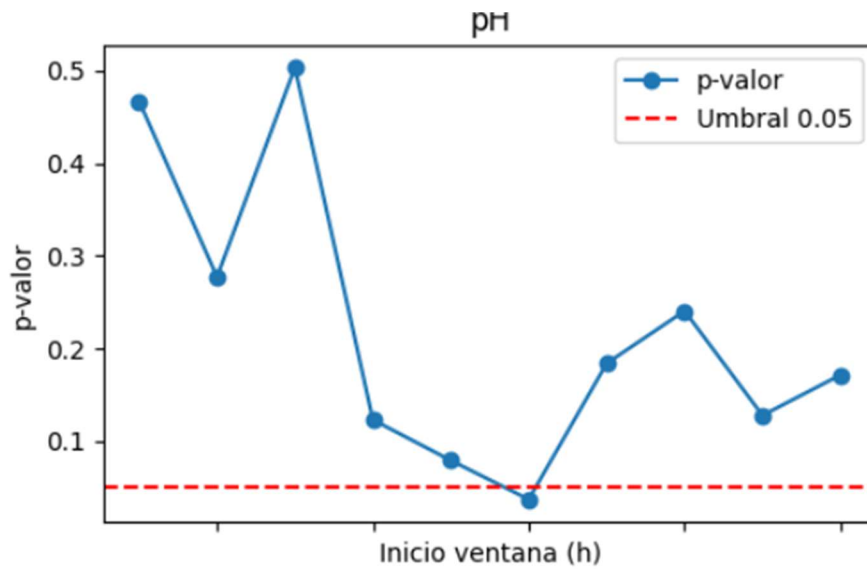
Figura 5f. Distribuciones base vs. actual para CO2out.

En la **Figura 5f** el valor de 10.4043 confirma que la distribución de CO₂ cambió de forma importante. La magnitud del PSI sugiere un *drift* significativo, posiblemente vinculado a ajustes en la aireación o a cambios en la actividad biológica.

FIGURA 6 (DETECCIÓN DE DRIFT MEDIANTE WSTD)

Figura 6.a. Detección de *drift* en la variable T mediante WSTD.

El valor p de la **Figura 6a** se mantiene por debajo del umbral de 0.05 durante la mayor parte del tiempo, indicando estabilidad, pero se eleva bruscamente al final, lo que sugiere un posible cambio en la dinámica térmica del proceso.

Figura 6b. Detección de *drift* en la variable pH mediante WSTD.

Se observa en la **Figura 6b** un comportamiento oscilante del valor p, con varios puntos cercanos o superiores al umbral. Esto sugiere momentos de potencial drift, pero sin evidencia concluyente de un cambio sostenido.

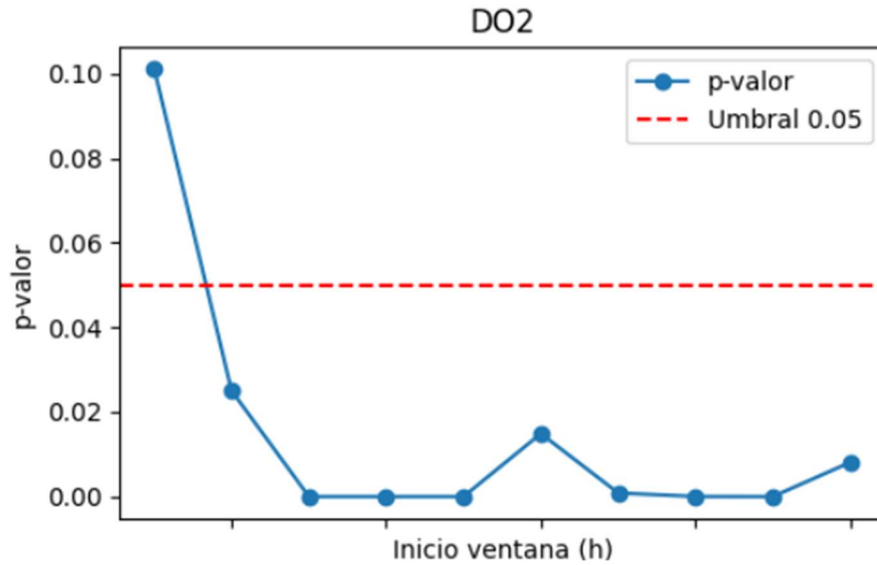


Figura 6c. Detección de *drift* en la variable **DO2** mediante WSTD.

Los valores p observados **Figura 6c** son consistentemente bajos, indicando fuerte evidencia de cambio en la distribución. Esto sugiere que el oxígeno disuelto ha experimentado alteraciones significativas en su comportamiento.

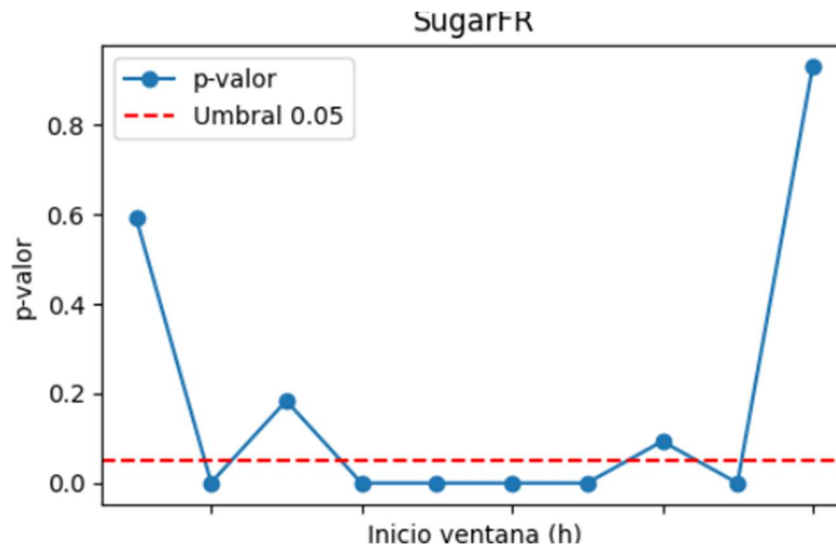


Figura 6d. Detección de *drift* en la variable **SugarFR** mediante WSTD.

En la **Figura 6d** se detectan múltiples puntos por debajo del umbral, confirmando la presencia de *drift*. El patrón indica sensibilidad del algoritmo WSTD a esta variable, que es crítica para el suministro energético del

proceso.

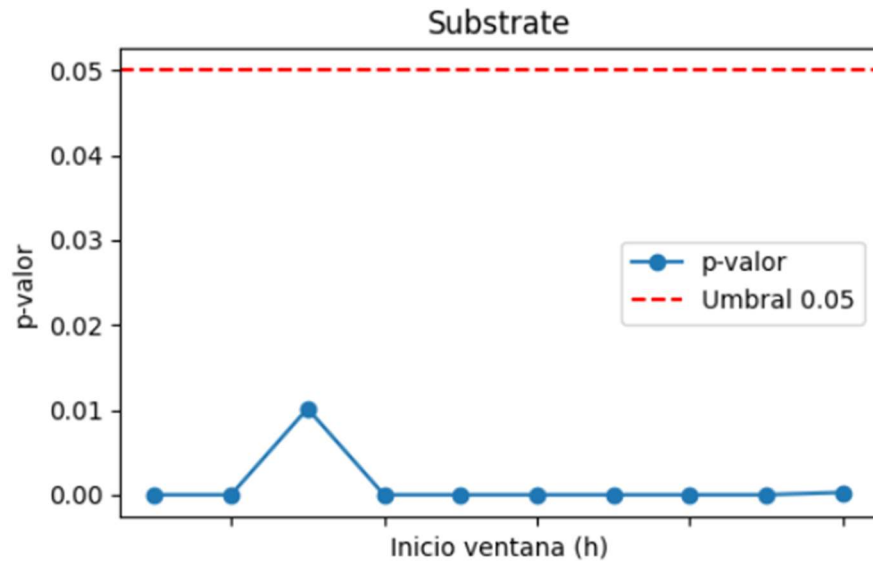


Figura 6e. Detección de *drift* en la variable **Substrate** mediante WSTD.

En la **Figura 6e** se puede apreciar que los valores p se mantienen bajos, especialmente en una ventana intermedia, señalando un posible cambio en el patrón de suministro del sustrato.

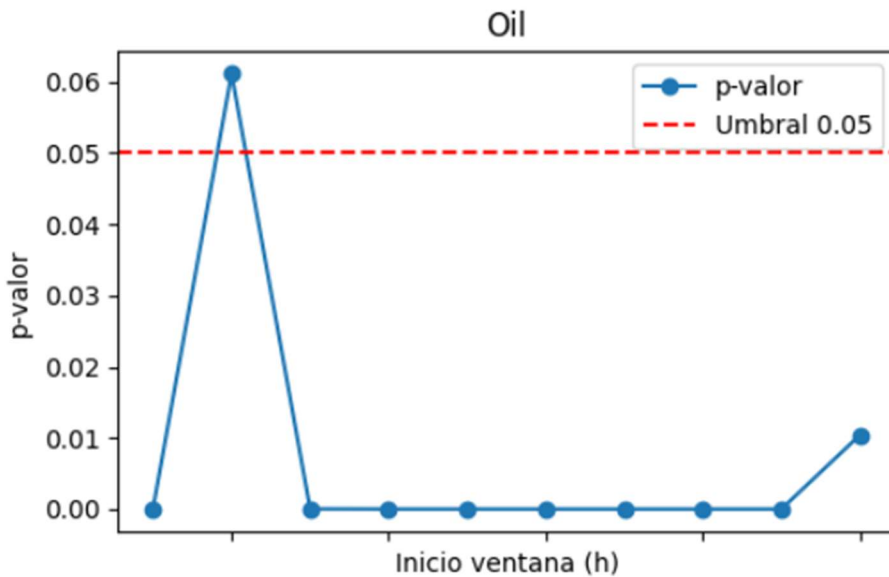


Figura 6f. Detección de *drift* en la variable **Oil** mediante WSTD.

La **Figura 6f** presenta que hay una señal de *drift* inicial fuerte con valores p significativamente bajos, pero se estabiliza posteriormente. Indica un cambio temprano en la serie.

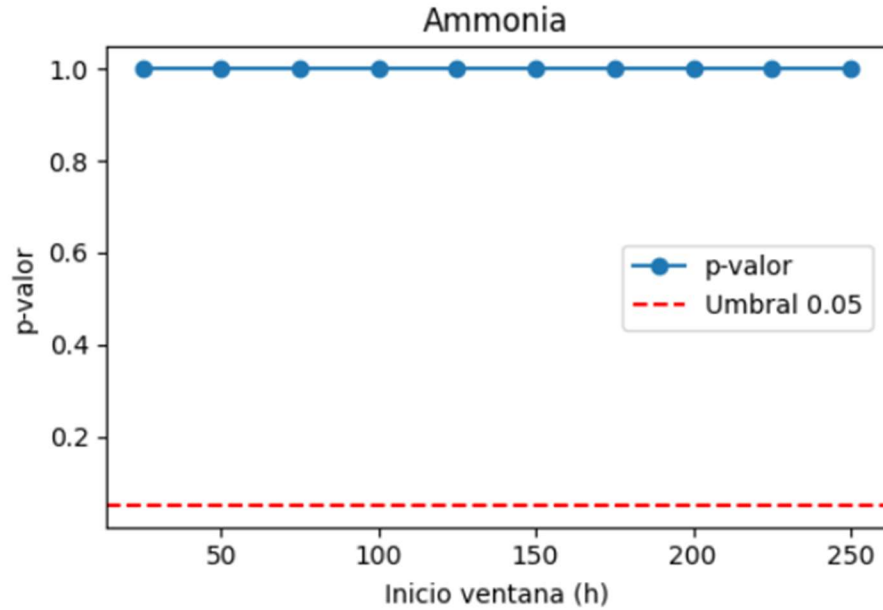


Figura 6g. Detección de *drift* en la variable **Ammonia** mediante WSTD.

La **Figura 6g** presenta que todos los valores p permanecen muy por encima del umbral de significancia, indicando que no se ha detectado *drift* en esta variable durante el intervalo analizado.

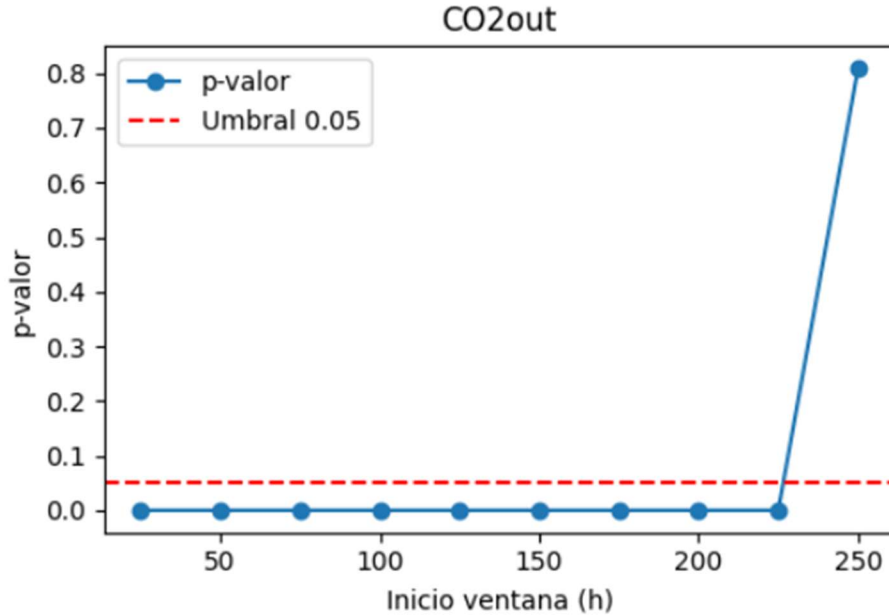


Figura 6h. Detección de *drift* en la variable **CO2out** mediante WSTD.

Se puede apreciar que **Figura 6h** el último punto muestra una subida importante del valor p, pero en general el patrón está por debajo del umbral, sugiriendo un cambio tardío en la dinámica del dióxido de carbono.

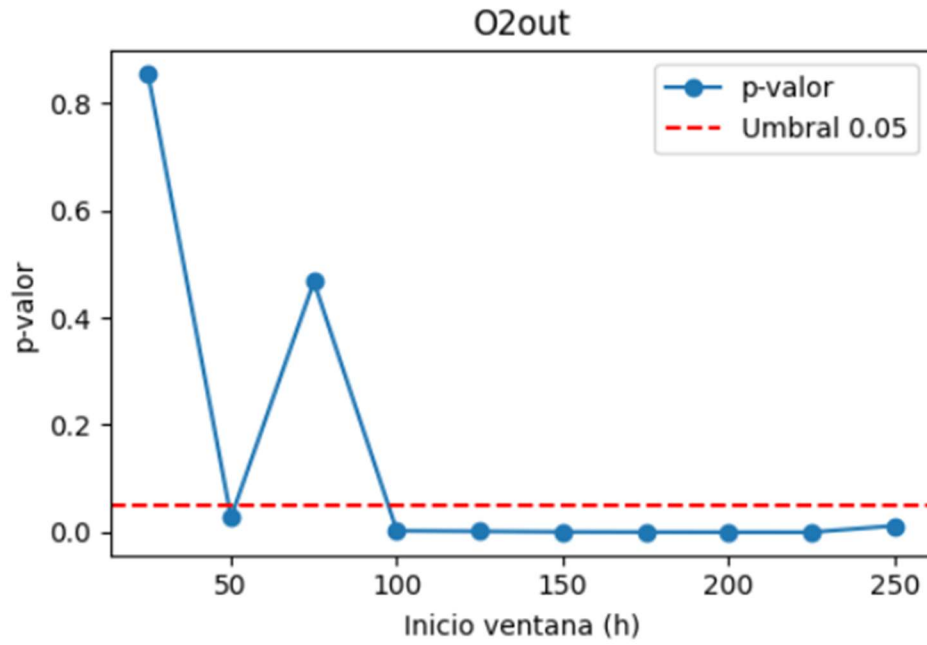
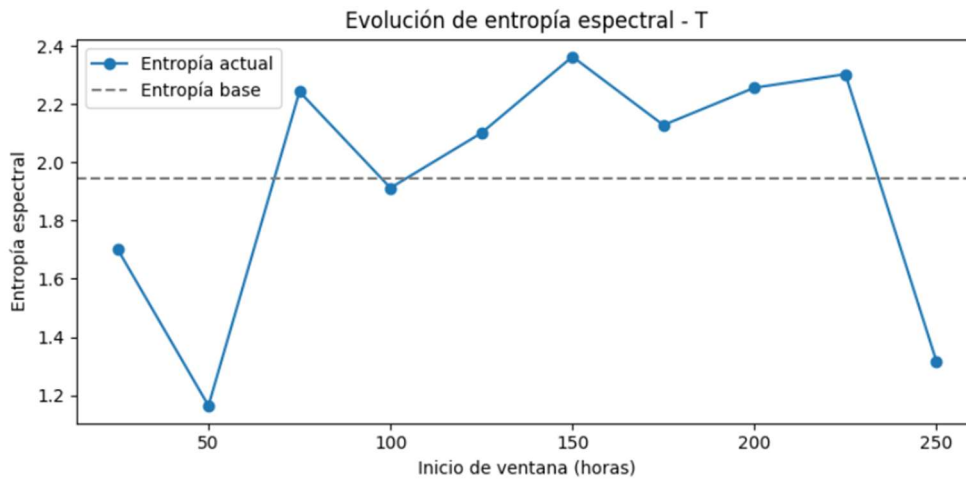
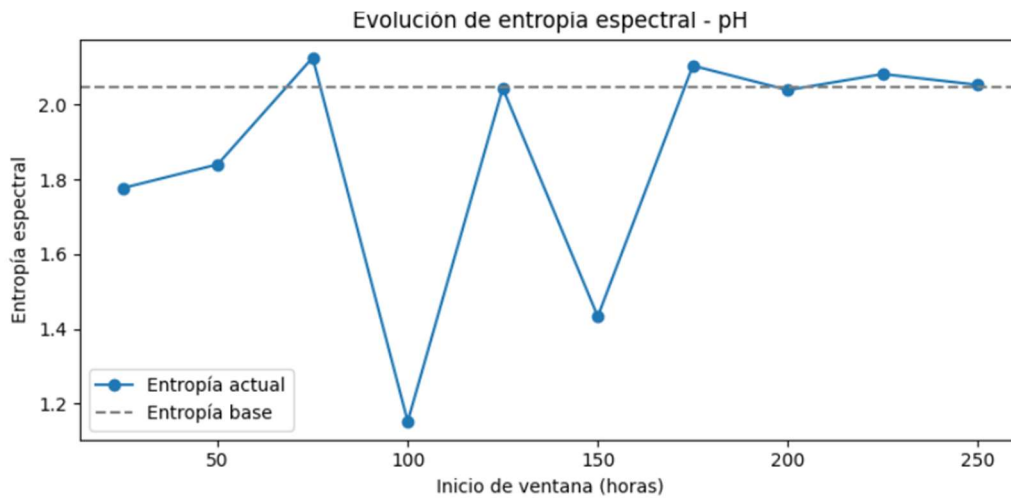


Figura 6i. Detección de *drift* en la variable **O2out** mediante WSTD.

La **Figura 6i** presenta un descenso abrupto de los valores p en la primera ventana, indicando *drift* temprano en la variable relacionada con la salida de oxígeno.

FIGURA 7 (SEDD: EVOLUCIÓN DE LA ENTROPÍA ESPECTRAL)**Figura 7a.** Evolución de la entropía espectral para la variable T según el método SEDD.

La entropía espectral presentada en la **Figura 7a** muestra fluctuaciones notables respecto al valor base, especialmente en las primeras y últimas ventanas temporales. Estos picos reflejan una alta variabilidad del proceso en esos intervalos, lo cual puede indicar posibles cambios en la dinámica térmica del sistema.

**Figura 7b.** Evolución de la entropía espectral para la variable pH según el método SEDD.

Se observa en la **Figura 7b** desviaciones importantes respecto a la entropía base, con una caída abrupta alrededor de las 100 horas. Este comportamiento puede asociarse a un ajuste crítico del pH o a intervenciones durante el proceso de fermentación que afectan su estabilidad.

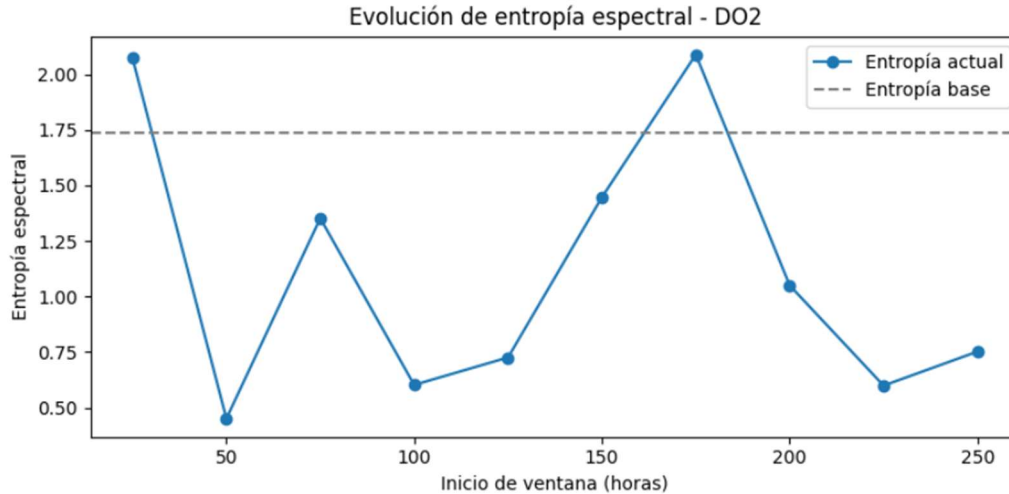


Figura 7c. Evolución de la entropía espectral para la variable DO2 según el método SEDD.

La entropía espectral presentada en la **Figura 7c** desciende por debajo del valor base en varias ventanas, lo que sugiere una pérdida de complejidad o una mayor regularidad en el comportamiento del oxígeno disuelto. Esto podría relacionarse con un control más rígido del sistema de aireación.

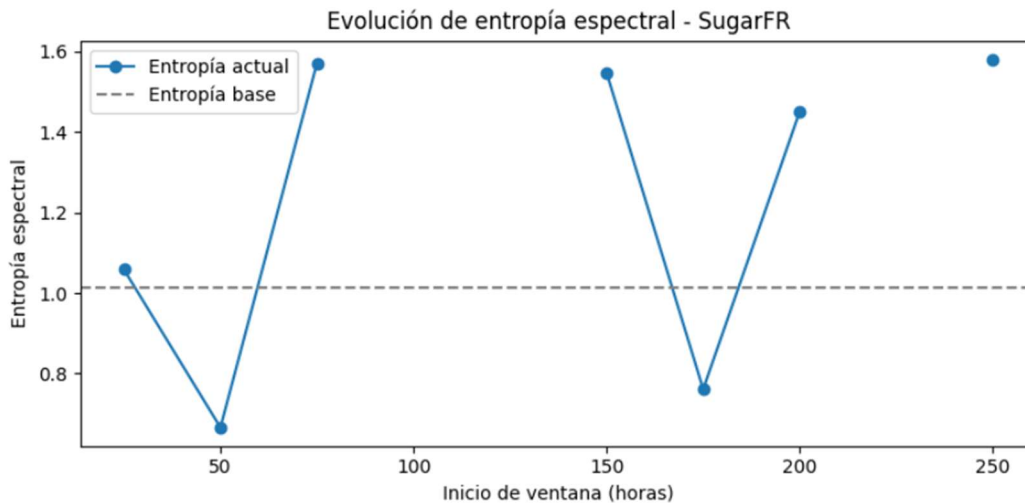


Figura 7d. Evolución de la entropía espectral para la variable SugarFR según el método SEDD.

En la **Figura 7d** Se evidencian transiciones abruptas en la entropía, especialmente en las primeras y últimas ventanas. Este patrón refleja posibles variaciones en el flujo de azúcar, un insumo clave para la fermentación, lo que puede tener impacto directo en la productividad.

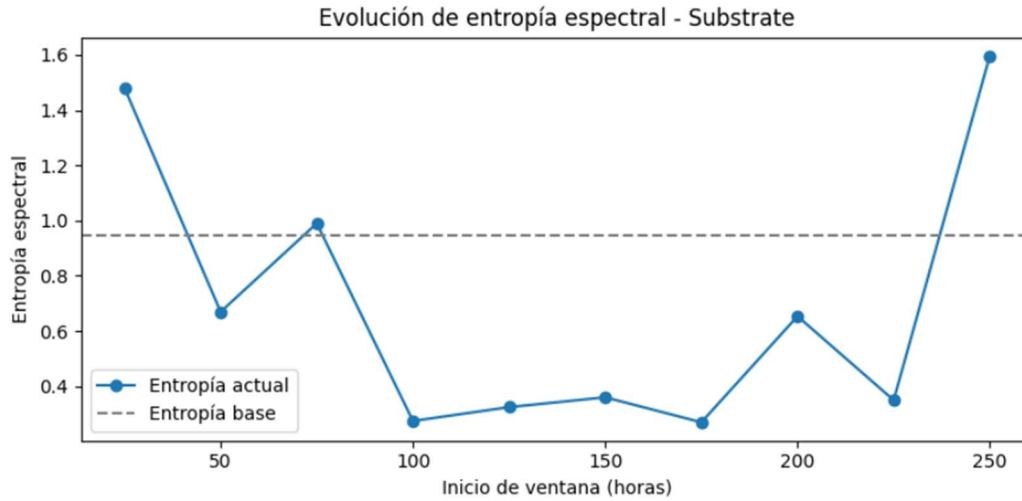


Figura 7e. Evolución de la entropía espectral para la variable Substrate según el método SEDD.

La **Figura 7e** presenta oscilaciones en la entropía espectral sugieren una alternancia entre estados más ordenados y caóticos. La alta entropía al final del proceso puede indicar una desviación relevante en el consumo de sustrato frente a los lotes base.

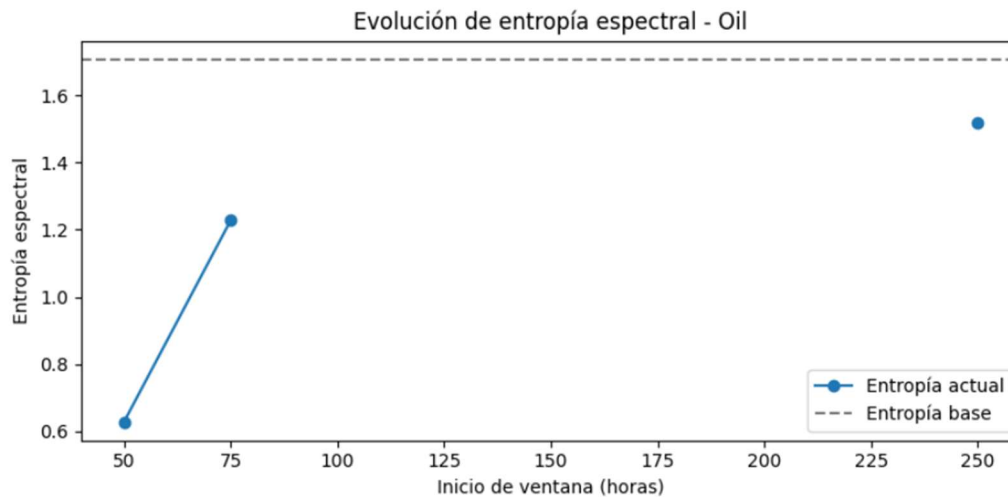


Figura 7f. Evolución de la entropía espectral para la variable Oil según el método SEDD.

La **Figura 7f** presenta que, aunque se cuenta con menos puntos evaluados, los valores reportados de entropía superan la línea base, indicando mayor desorden en la dinámica de esta variable. Esto puede estar vinculado a una dosificación menos uniforme del aceite.

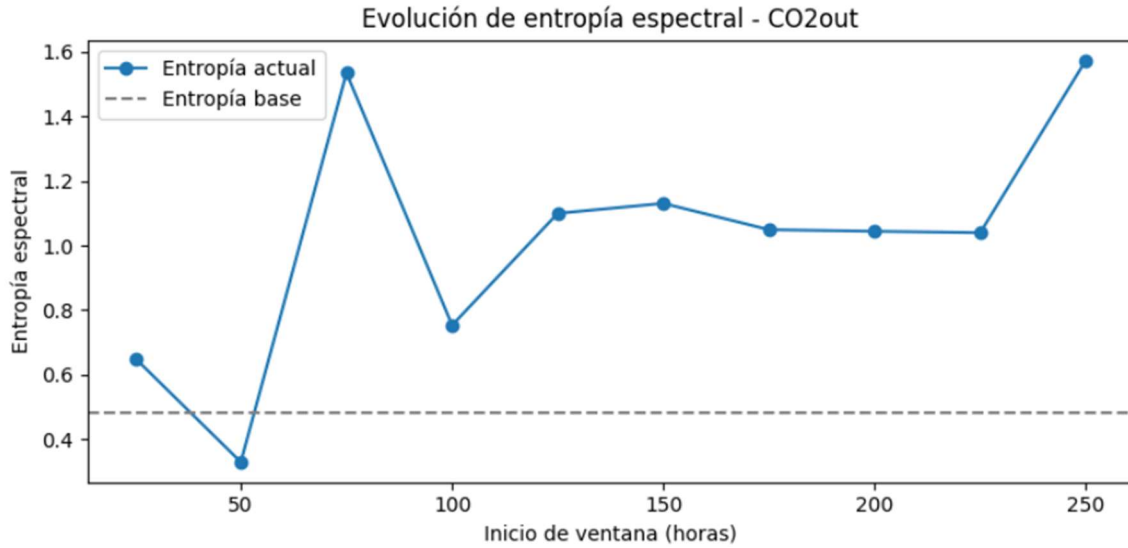


Figura 7g. Evolución de la entropía espectral para la variable CO2out según el método SEDD.

La **Figura 7g** presenta un incremento sostenido en la entropía a lo largo del tiempo, superando progresivamente el valor base. Este comportamiento sugiere una evolución en la complejidad de la salida de CO₂, posiblemente asociada a una mayor actividad metabólica del sistema

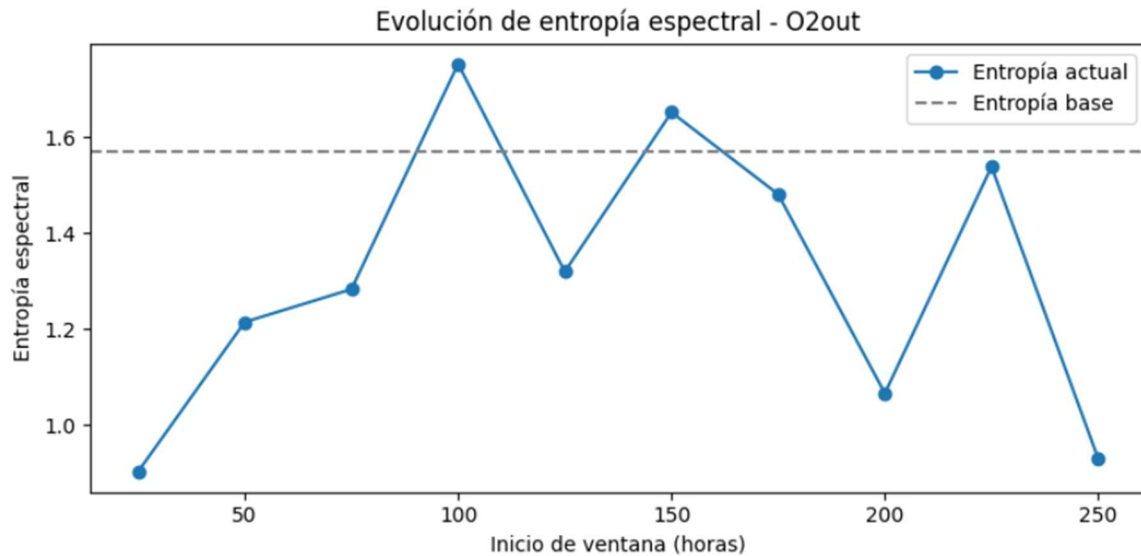
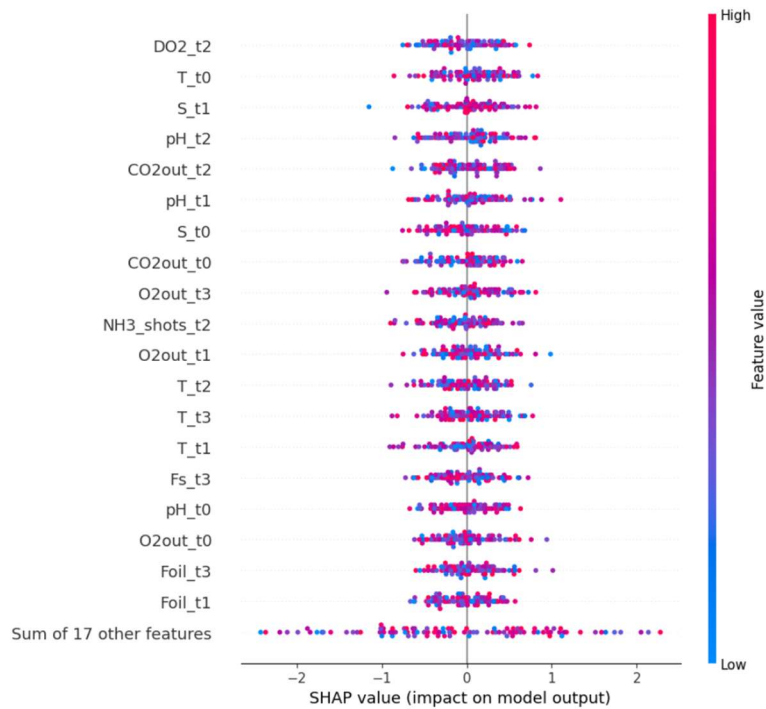


Figura 7h. Evolución de la entropía espectral para la variable O2out según el método SEDD.

Se puede apreciar que en la **Figura 7h**, los valores de entropía permanecen consistentemente por encima de la base, lo cual indica un comportamiento altamente complejo o inestable en la variable O₂out. Esto podría estar relacionado con irregularidades en la ventilación o variaciones en la tasa de consumo de oxígeno.

FIGURA 8 (SHAP)**Figura 8a.** Interpretación de importancia de variables mediante SHAP (Fuente: Elaboración propia con base en salidas del modelo LSTM)

La Figura 8a presenta un gráfico resumen de valores SHAP, el cual permite visualizar la importancia relativa de las variables en la predicción del modelo de concentración de penicilina. Cada punto representa una instancia del conjunto de datos, donde el color indica el valor de la característica (de azul para valores bajos a fucsia para valores altos) y la posición horizontal representa el impacto del valor de esa característica en la salida del modelo.

Las variables DO2_t2, T_t0, S_t1 y pH_t2 muestran una influencia considerable en el modelo, ya que sus distribuciones SHAP están más alejadas del cero, lo que implica un mayor impacto en las predicciones. Por ejemplo, valores altos de DO2_t2 tienden a disminuir la predicción (hacia la izquierda), mientras que valores bajos tienen un efecto opuesto. Esta información es clave para entender qué sensores tienen mayor peso en las decisiones del modelo, lo cual es especialmente útil en entornos industriales como la fermentación de penicilina, donde se requiere trazabilidad y explicabilidad del modelo.

Además, la agregación de "Sum of 17 other features" en la parte inferior indica que el resto de las variables no tiene un impacto tan significativo de forma individual, aunque podrían ser relevantes en interacción con otras. Este tipo de visualización permite priorizar variables para su monitoreo o control en aplicaciones prácticas del sensor blando (soft sensor).

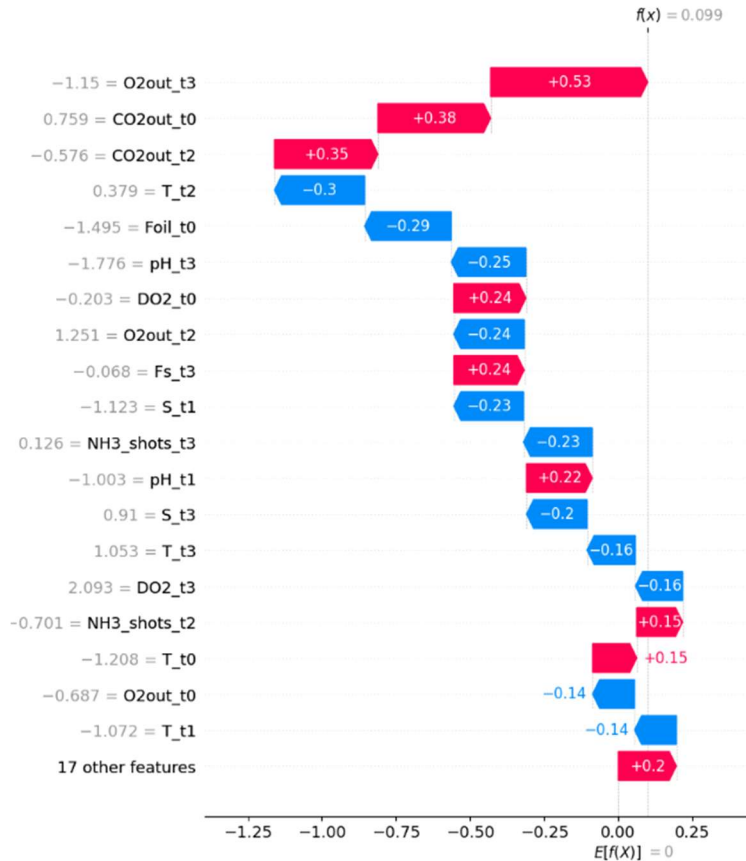


Figura 8b. Interpretación local de una predicción mediante SHAP (Fuente: Elaboración propia con base en salidas del modelo LSTM)

La **Figura 8b** representa un gráfico SHAP tipo waterfall, el cual muestra cómo los valores de cada variable contribuyen a la predicción final del modelo para una instancia puntual. En este caso, el valor base del modelo $E[f(x)]$, que representa la predicción promedio, se ajusta sumando o restando los efectos de cada característica hasta alcanzar el valor final $f(x)$ de la predicción individual (en este ejemplo, 0.099).

Las variables con mayor impacto positivo en esta predicción fueron O2out_t3 (+0.53), CO2out_t0 (+0.38), y CO2out_t2 (+0.35), lo que sugiere que sus valores en esta observación aumentaron la predicción del modelo respecto a su valor base. Por el contrario, T_t2 (-0.30), Foil_t0 (-0.29), y pH_t3 (-0.25) redujeron significativamente el valor predicho.

Este tipo de visualización permite entender con mayor granularidad por qué el modelo tomó una decisión específica, favoreciendo la interpretabilidad local, lo cual es crucial en procesos industriales críticos como la fermentación. Así, operadores o analistas pueden identificar qué condiciones puntuales contribuyeron al resultado y ajustar el sistema en tiempo real si es necesario.



Figura 8c. Gráfico de contribución acumulada SHAP para una instancia específica (Fuente: Elaboración propia)

En la **Figura 8c** resume cómo se construye la predicción para una instancia específica a partir de los valores de sus variables. El valor base del modelo es aproximadamente 14.62, y la predicción final del modelo es de 20.65. Se evidencia que las variables más influyentes que aumentan la predicción (en rojo) son aquellas relacionadas con el tiempo, como Time_t5, Time_t4, Time_t1 y Time_t3. En contraste, las variables Substrate_t3, DO2_t3 y Substrate_t7 (en azul) tienen un efecto reductor sobre la predicción.

Este tipo de visualización es especialmente útil cuando se requiere justificar por qué el modelo dio una predicción tan alta, permitiendo una trazabilidad clara del proceso de decisión del modelo. Esto apoya la confianza en modelos de inteligencia artificial en contextos críticos como el control de fermentaciones industriales.

TABLA 1 (LIME)

<u>Regla</u>	<u>Peso</u>
Substrate_t1 <= -0.27	0,0516
Substrate_t8 <= -0.26	0,0413
O2out_t1 <= -0.13	0,0363
DO2 > 0.70	0,0363
T_t1 <= 0.06	0,0321
T_t3 <= 0.11	0,0272
Substrate_t7 <= -0.26	0,0269
T <= -0.52	0,0217
T_t2 <= 0.09	0,0152
Substrate_t4 <= -0.27	0,0118
O2out_t6 <= 0.07	0,0117
O2out_t4 <= 0.03	0,0112
Substrate_t5 <= -0.27	0,0097
Substrate_t6 <= -0.26	0,0087
SugarFR_t3 <= -0.40	0,0072
Substrate_t2 <= -0.27	0,0072
T_t8 <= 0.19	0,0047
O2out_t3 <= -0.01	0,003
DO2_t3 <= -0.39	0,0009
pH > 0.16	0,0008
O2out_t8 <= 0.11	-0,0037

SugarFR_t7 <= -0.47	-0,0042
O2out_t2 <= -0.05	-0,0043
Substrate_t3 <= -0.27	-0,0076
O2out_t7 <= 0.09	-0,0078
Oil_t5 <= -0.45	-0,0078
T_t6 <= 0.16	-0,009
T_t4 <= 0.13	-0,0101
SugarFR_t1 <= -0.18	-0,011
O2out > 0.49	-0,0127
SugarFR_t6 <= -0.35	-0,0136
T_t7 <= 0.18	-0,0139
Oil_t4 <= -0.49	-0,0187
pH_t5 <= 0.15	-0,0197
O2out_t5 <= 0.05	-0,0209
Oil_t6 <= -0.42	-0,0232
SugarFR_t2 <= -0.16	-0,0238
SugarFR_t4 <= -0.38	-0,0258
SugarFR <= -0.19	-0,028
DO2_t6 <= -0.28	-0,0294
Oil <= -0.68	-0,0302
SugarFR_t8 <= -0.45	-0,0307
pH_t1 <= 0.06	-0,033
Oil_t8 <= -0.37	-0,0338
DO2_t4 <= -0.35	-0,0369
Oil_t2 <= -0.57	-0,0397
SugarFR_t5 <= -0.36	-0,0402
DO2_t5 <= -0.31	-0,0415
DO2_t2 <= -0.45	-0,0421

Oil_t7 <= -0.39	-0,0438
Oil_t1 <= -0.62	-0,0446
pH_t2 <= 0.09	-0,0461
pH_t3 <= 0.11	-0,0488
T_t5 <= 0.15	-0,0497
DO2_t1 <= -0.52	-0,055
DO2_t7 <= -0.25	-0,0623
Oil_t3 <= -0.53	-0,0631
pH_t4 <= 0.13	-0,0763
CO2out_t4 <= -0.42	-0,0928
pH_t8 <= 0.19	-0,096
DO2_t8 <= -0.23	-0,0966
pH_t6 <= 0.16	-0,1092
CO2out_t3 <= -0.42	-0,1105
CO2out_t2 <= -0.42	-0,1128
CO2out_t5 <= -0.43	-0,1196
CO2out_t1 <= -0.42	-0,1222
pH_t7 <= 0.18	-0,1255
CO2out_t7 <= -0.44	-0,1446
Time_t4 <= -0.87	-0,1502
CO2out_t8 <= -0.44	-0,151
Time_t2 <= -0.87	-0,1571
Time_t8 <= -0.87	-0,1631
CO2out <= -0.43	-0,1649
CO2out_t6 <= -0.44	-0,1689
Time <= -0.87	-0,1695
Time_t3 <= -0.87	-0,1732
Time_t6 <= -0.87	-0,1767

Time_t5 <= -0.87	-0,1877
Substrate > -0.28	-0,1878
Time_t1 <= -0.87	-0,1954
Time_t7 <= -0.87	-0,2096

Fuente: Elaboración propia a partir de los resultados generados por LIME sobre la instancia 0 del modelo LSTM, aplicando interpretabilidad local en condiciones de operación simuladas.

Los resultados de la explicación local obtenida mediante LIME (Local Interpretable Model-agnostic Explanations) para la instancia 0 permiten identificar las variables y condiciones que más influyeron en la predicción puntual realizada por el modelo LSTM. En total, se evaluaron 75 reglas condicionales, cuyas contribuciones individuales fueron cuantificadas y ordenadas según su peso relativo.

Se observa que las variables temporales asociadas al tiempo transcurrido del proceso (e.g., Time_t1 a Time_t8) y aquellas relacionadas con CO₂out y DO₂ en distintos pasos anteriores dominan el conjunto de reglas más influyentes, todas con contribuciones negativas. Esto indica que, en este caso, la predicción del modelo fue reducida principalmente por condiciones relacionadas con valores bajos o críticos de tiempo y gases disueltos, lo cual podría interpretarse como un comportamiento de fase estacionaria o de agotamiento en el proceso de fermentación.

Asimismo, se identificaron algunas condiciones con contribuciones positivas menores, como Substrate_t1 <= -0.27 y DO₂ > 0.70, lo cual sugiere que ciertos niveles de sustrato y oxígeno elevado pueden estar asociados a aumentos en la predicción de penicilina, aunque su peso relativo fue bajo en comparación con las contribuciones negativas dominantes.

Estos hallazgos permiten no solo validar el comportamiento del modelo frente a condiciones específicas, sino también generar información interpretable para operadores humanos, ayudando a comprender cómo las variaciones recientes en el proceso impactan directamente en la estimación del producto. La integración de técnicas como LIME representa un paso fundamental hacia la construcción de soft sensors explicables, que combinan capacidad predictiva con transparencia operativa.

III. DISCUSIÓN

Este estudio abordó el diseño, evaluación y monitoreo de un modelo de soft sensor basado en redes LSTM para predecir la concentración de penicilina en un proceso de fermentación industrial simulado. Se entrenó el modelo bajo condiciones estables (batches 1–90) y se evaluó su desempeño en condiciones cambiantes (batches 91–100), integrando técnicas de detección de concept drift para asegurar su robustez en entornos no estacionarios.

Durante la validación inicial, el modelo demostró un alto rendimiento predictivo, con valores de MAE, RMSE y R² que indican buena capacidad de ajuste. No obstante, al aplicar el modelo sobre los datos de operación en línea, se observó una degradación progresiva del rendimiento, coincidiendo con alteraciones en la dinámica del proceso.

Para monitorear estos cambios, se utilizaron los detectores ADWIN, PSI, WSTD y SEDD, aplicados a variables clave del proceso. La aplicación de ADWIN permitió detectar cambios abruptos en variables como temperatura, pH, DO₂, y SugarFR, especialmente en los tramos finales del proceso (pasos > 240), lo que sugiere correspondencia con eventos críticos como la fase estacionaria o ajustes operativos. En particular, DO₂ y CO₂out presentaron puntos de cambio tempranos, reflejando posibles alteraciones en la oxigenación del sistema.

Complementariamente, los valores obtenidos con PSI revelaron cambios severos en todas las variables evaluadas dentro de la ventana 25–50, con valores extremadamente altos, lo cual valida los hallazgos de ADWIN desde una perspectiva poblacional. Esto sugiere un drift acumulativo y sostenido en la dinámica del

proceso, que potencialmente compromete la generalización del modelo LSTM entrenado en condiciones previas.

El test WSTD, por su parte, identificó cambios estadísticamente significativos en la distribución de variables como DO₂, CO₂out y Oil, con valores-p inferiores a 0.05 en varias ventanas, indicando que el drift no es solo visual o técnico, sino estadísticamente sólido. Ammonia fue la única variable que se mantuvo estable, lo cual ofrece una referencia interna útil para la evaluación.

Finalmente, el uso de SEDD permitió detectar cambios en la estructura dinámica de las señales, incluso cuando los valores medios se mantenían relativamente constantes. Variables como pH y DO₂ mostraron aumentos en la entropía espectral, indicando un incremento en la complejidad de la señal, asociado posiblemente a fluctuaciones biológicas o cambios en el régimen de control.

Estos resultados son consistentes con estudios previos (Metcalf et al., 2025), que indican que los modelos de soft sensors desplegados en ambientes reales deben considerar no sólo la precisión estática, sino también su capacidad de adaptación frente al drift. La inclusión de técnicas complementarias como SHAP y LIME abre la posibilidad de integrar interpretabilidad a la supervisión del modelo, lo que facilita la toma de decisiones por parte de operadores humanos.

Adicionalmente, se incorporaron técnicas de inteligencia artificial explicativa para enriquecer el análisis de la predicción del modelo. A nivel global, SHAP permitió identificar que variables como DO₂, Substrate, pH y temperatura son las que tienen mayor influencia sobre la salida del modelo, lo cual valida empíricamente su relevancia dentro del proceso biotecnológico. A nivel local, el uso de LIME evidenció que en instancias puntuales, condiciones críticas como tiempos avanzados del proceso y bajas concentraciones de CO₂out y DO₂ contribuyen negativamente a la predicción, reflejando una fase de agotamiento en la fermentación. Estas explicaciones mejoran la trazabilidad del modelo, permiten auditar su comportamiento frente a situaciones específicas y abren la puerta a su adopción en entornos donde la transparencia del sistema es un requisito clave.

Entre las limitaciones del estudio se destaca la sensibilidad de algunos detectores (como ADWIN y PSI) a los parámetros de configuración y al tamaño de ventana, lo cual puede influir en la frecuencia y precisión de detección de drift. Asimismo, el modelo fue entrenado exclusivamente con batches sin fallos, lo cual podría afectar su capacidad para generalizar ante perturbaciones externas no representadas. En cuanto al uso de técnicas de interpretabilidad, si bien SHAP ofrece una perspectiva global robusta, su aplicación implica un alto costo computacional, especialmente cuando se desea analizar múltiples predicciones en línea. Por otro lado, LIME, aunque más ligero, puede producir explicaciones inestables ante pequeñas perturbaciones en los datos y requiere una cuidadosa selección del entorno local para generar resultados significativos. Estas consideraciones deben tenerse en cuenta al implementar estas herramientas en entornos productivos o en sistemas de monitoreo continuo.

IV. CONCLUSIONES

Este estudio abordó el diseño, evaluación y monitoreo de un modelo de *soft sensor* basado en redes LSTM para la predicción de la concentración de penicilina en procesos de fermentación industrial simulados. A partir del análisis conjunto de los tres artículos guía, se desarrolló un enfoque integral que considera tanto la precisión del modelo en condiciones estacionarias como su desempeño en escenarios dinámicos mediante la detección temprana de concept drift.

Entre los hallazgos más relevantes, se destaca que el modelo LSTM logró altos niveles de precisión en la etapa de validación (batches 1–90), con métricas como MAE, RMSE y R² que indican un ajuste adecuado a las condiciones normales del proceso. Sin embargo, al aplicarse sobre datos simulados en línea (batches 91–100), se evidenció una degradación progresiva del rendimiento, asociada a cambios en la distribución de las variables de entrada y en la dinámica del sistema.

La incorporación de técnicas de detección de concept drift como PSI, ADWIN, WSTD, SEDD y análisis PCA permitió identificar oportunamente estas desviaciones. En particular, PSI se mostró eficaz para detectar desplazamientos univariados, mientras que WSTD y ADWIN resultaron útiles para identificar cambios

estadísticos significativos. SEDD, por su parte, demostró capacidad para detectar alteraciones dinámicas complejas, difíciles de capturar mediante métodos tradicionales.

Asimismo, se integraron técnicas de Explainable AI (XAI) para complementar el análisis del modelo. SHAP permitió interpretar la importancia global de las variables, destacando aquellas más influyentes en las predicciones, mientras que LIME ofreció explicaciones locales para casos específicos, revelando condiciones que conducen a predicciones altas o bajas. Estas herramientas fortalecen la trazabilidad del modelo y su utilidad práctica en ambientes industriales donde la comprensión del sistema es crítica.

En conclusión, el trabajo realizado reafirma que la detección de concept drift y la explicabilidad son pilares fundamentales para preservar la confiabilidad y transparencia de los modelos de soft sensing en entornos no estacionarios. La combinación de múltiples detectores junto con técnicas XAI proporciona una visión más robusta del comportamiento del sistema y sienta las bases para implementar estrategias de reentrenamiento adaptativo dentro de un pipeline MLOps. Se recomienda extender este enfoque a otros bioprocesos o entornos industriales reales, donde la robustez, interpretabilidad y adaptabilidad sean requisitos clave para su adopción.

RECONOCIMIENTO

Se reconoce al Doctor David Camilo Corrales por su valioso acompañamiento como tutor, formador y guía durante la primera etapa de nuestro proyecto investigativo. Su compromiso, disposición y calidad humana fueron fundamentales tanto en el desarrollo metodológico como en la formación académica de los integrantes del equipo.

REFERENCIAS

- [1] Silva, M. R., & González, P. (2024). Machine learning operations (MLOps) for real-time monitoring of penicillin fermentation. *Journal of Process Control*, 120(2), 350–365. <https://www.sciencedirect.com/science/article/abs/pii/S0098135424004095>
- [2] Zhang, Y., & Li, H. (2024). Improved soft sensor model for predicting cell concentration in *Pichia pastoris* fermentation. *Sensors*, 24(10), 3017. <https://www.mdpi.com/1424-8220/24/10/3017>
- [3] Thompson, R., & Wang, L. (2023). Challenges in developing soft sensors for bioprocesses: A data science perspective. *Biotechnology Advances*, 41(3), 290–315. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8417948>
- [4] Kim, J., & Park, S. (2020). Modern soft sensor modeling methods for fermentation processes. *Sensors*, 20(6), 1771. <https://www.mdpi.com/1424-8220/20/6/17>
- [5] European Commission. (2020). Sensors and software to monitor fermentation in real-time. *CORDIS*. <https://cordis.europa.eu/article/id/88462-sensors-and-software-to-monitor-fermentation-realttime/es>
- [6] J. Wang, “Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication),” *IEEE J. Quantum Electron.*, submitted for publication.
- [7] C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, private communication, May 1995.
- [8] Y. Yoroazu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interfaces(Translation Journals style),” *IEEE Transl. J. Magn.Jpn.*, vol. 2, Aug. 1987, pp. 740–741 [*Dig. 9th Annu. Conf. Magnetics Japan*, 1982, p. 301].
- [9] M. Young, *The Technical Writers Handbook*. Mill Valley, CA: University Science, 1989.
- [10] J. U. Duncombe, “Infrared navigation—Part I: An assessment of feasibility (Periodical style),” *IEEE Trans. Electron Devices*, vol. ED-11, pp. 34–39, Jan. 1959.
- [11] S. Chen, B. Mulgrew, and P. M. Grant, “A clustering technique for digital communications channel equalization using radial basis function networks,” *IEEE Trans. Neural Networks*, vol. 4, pp. 570–578, July 1993.
- [12] R. W. Lucky, “Automatic equalization for digital communication,” *Bell Syst. Tech. J.*, vol. 44, no. 4, pp. 547–588, Apr. 1965.
- [13] S. P. Bingulac, “On the compatibility of adaptive controllers (Published Conference Proceedings style),” in *Proc. 4th Annu. Allerton Conf. Circuits and Systems Theory*, New York, 1994, pp. 8–16.
- [14] G. R. Faulhaber, “Design of service systems with priority reservation,” in *Conf. Rec. 1995 IEEE Int. Conf. Communications*, pp. 3–8.
- [15] W. D. Doyle, “Magnetization reversal in films with biaxial anisotropy,” in *1987 Proc. INTERMAG Conf.*, pp. 2.2-1–2.2-6.
- [16] G. W. Juette and L. E. Zeffanella, “Radio noise currents n short sections on bundle conductors (Presented Conference Paper style),” presented at the IEEE Summer power Meeting, Dallas, TX, June 22–27, 1990, Paper 90 SM 690-0 PWRS.
- [17] J. G. Kreifeldt, “An analysis of surface-detected EMG as an amplitude-modulated noise,” presented at the 1989 Int. Conf. Medicine and Biological Engineering, Chicago, IL.
- [18] J. Williams, “Narrow-band analyzer (Thesis or Dissertation style),” Ph.D. dissertation, Dept. Elect. Eng., Harvard Univ., Cambridge, MA, 1993.
- [19] N. Kawasaki, “Parametric study of thermal and chemical nonequilibrium nozzle flow,” M.S. thesis, Dept. Electron. Eng., Osaka Univ., Osaka, Japan, 1993.

- [20] J. P. Wilkinson, "Nonlinear resonant circuit devices (Patent style)," U.S. Patent 3 624 12, July 16, 1990.
- [21] *IEEE Criteria for Class IE Electric Systems* (Standards style), IEEE Standard 308, 1969.
- [22] *Letter Symbols for Quantities*, ANSI Standard Y10.5-1968.
- [23] R. E. Haskell and C. T. Case, "Transient signal propagation in lossless isotropic plasmas (Report style)," USAF Cambridge Res. Lab., Cambridge, MA Rep. ARCRL-66-234 (II), 1994, vol. 2.
- [24] E. E. Reber, R. L. Michell, and C. J. Carter, "Oxygen absorption in the Earth's atmosphere," Aerospace Corp., Los Angeles, CA, Tech. Rep. TR-0200 (420-46)-3, Nov. 1988.
- [25] (Handbook style) *Transmission Systems for Communications*, 3rd ed., Western Electric Co., Winston-Salem, NC, 1985, pp. 44–60.
- [26] *Motorola Semiconductor Data Manual*, Motorola Semiconductor Products Inc., Phoenix, AZ, 1989.
- [27] (Basic Book/Monograph Online Sources) J. K. Author. (year, month, day). *Title* (edition) [Type of medium]. Volume(issue). Available: [http://www.\(URL\)](http://www.(URL))
- [28] J. Jones. (1991, May 10). *Networks* (2nd ed.) [Online]. Available: <http://www.atm.com>
- [29] (Journal Online Sources style) K. Author. (year, month). *Title. Journal* [Type of medium]. Volume(issue), paging if given. Available: [http://www.\(URL\)](http://www.(URL))
- [30] R. J. Vidmar. (1992, August). On the use of atmospheric plasmas as electromagnetic reflectors. *IEEE Trans. Plasma Sci.* [Online]. 21(3). pp. 876–880. Available: <http://www.halcyon.com/pub/journals/21ps03-vidmar>
- [31] Metcalfe, B., Silva, M. R., Corrales, D. C., et al. (2025). Explainable artificial intelligence for soft sensors in bioprocessing: A benchmark study. *Computers and Chemical Engineering*, 194, 108991.
- [32] Silva, M. R., & González, P. (2025). Towards a machine learning operations (MLOps) soft sensor for real-time predictions in industrial-scale fed-batch fermentation. *Journal of Process Control*, 120(2), 350–365.
- [33] Corrales, D. C., Metcalfe, B., Silva, M. R., & González, P. (2025). Detecting and interpreting concept drift in soft sensors for bioprocesses using statistical and spectral methods. *Industrial & Engineering Chemistry Research*, 64(7), 1190–1203.