



DIAGNÓSTICO DE FALLAS EN LA SOPLADORA DE ELECTROLIT: UN ENFOQUE BASADO EN ANÁLISIS EXPLORATORIO

Marissa E. Luna, Ximena A. Cantón, Nubia G. Barajas and Mariana L. Maldonado

CONTEXTO DEL PROBLEMA

En el proceso de fabricación de botellas de Electrolit, la sopladora es una máquina clave.

Su vida útil esperada es de 12 meses, pero actualmente falla en un rango de 2-3 meses, generando pérdidas económicas.

Se busca entender las posibles causas de estas fallas mediante el Análisis Exploratorio de Datos (EDA).

DESCRIPCIÓN DE LOS DATOS

- Datos en formato Parquet de los meses de noviembre y enero.
- Al combinar los archivos se tuvo una base con 90,798,747 observaciones.
- Alta complejidad --> debido al volumen de información.
- Datos sobre temperatura, presión, consumo de energía, entre otros.
- Se inició con 28 variables, pero algunas fueron eliminadas por baja representatividad.

DESAFÍOS INICIALES

- Volumen de Datos: Difícil almacenamiento y procesamiento.
- Datos Faltantes: Algunas variables tenían menos del 0.1% de registros útiles.
- Desalineación Temporal: Necesidad de ajustar la zona horaria (-6 horas).

PREPROCESAMIENTO DE DATOS

Corrección temporal

- Se ajustaron los timestamps restando 6 horas para alinear los registros con el horario real de operación.

Selección y Limpieza de Variables

- Se seleccionaron 39% variables clave para el análisis.
- Se eliminaron 8 variables con datos insuficientes.
- Se descartaron variables sin variabilidad significativa (desviación estándar =0).

Optimización para análisis

- División en chunks de 100,000 registros para facilitar el procesamiento.
- Creación de archivos CSV con datos procesados para su análisis detallado.

SELECCIÓN Y LIMPIEZA DE VARIABLES

Partiendo de una muestra representativa del 39% del total de los datos.

1. Seleccionamos y eliminamos variables con menos del 0.1% de observaciones respecto al total:

- Baja representatividad.
- Posible falta de impacto.
- Ruido en el análisis sin un beneficio claro,

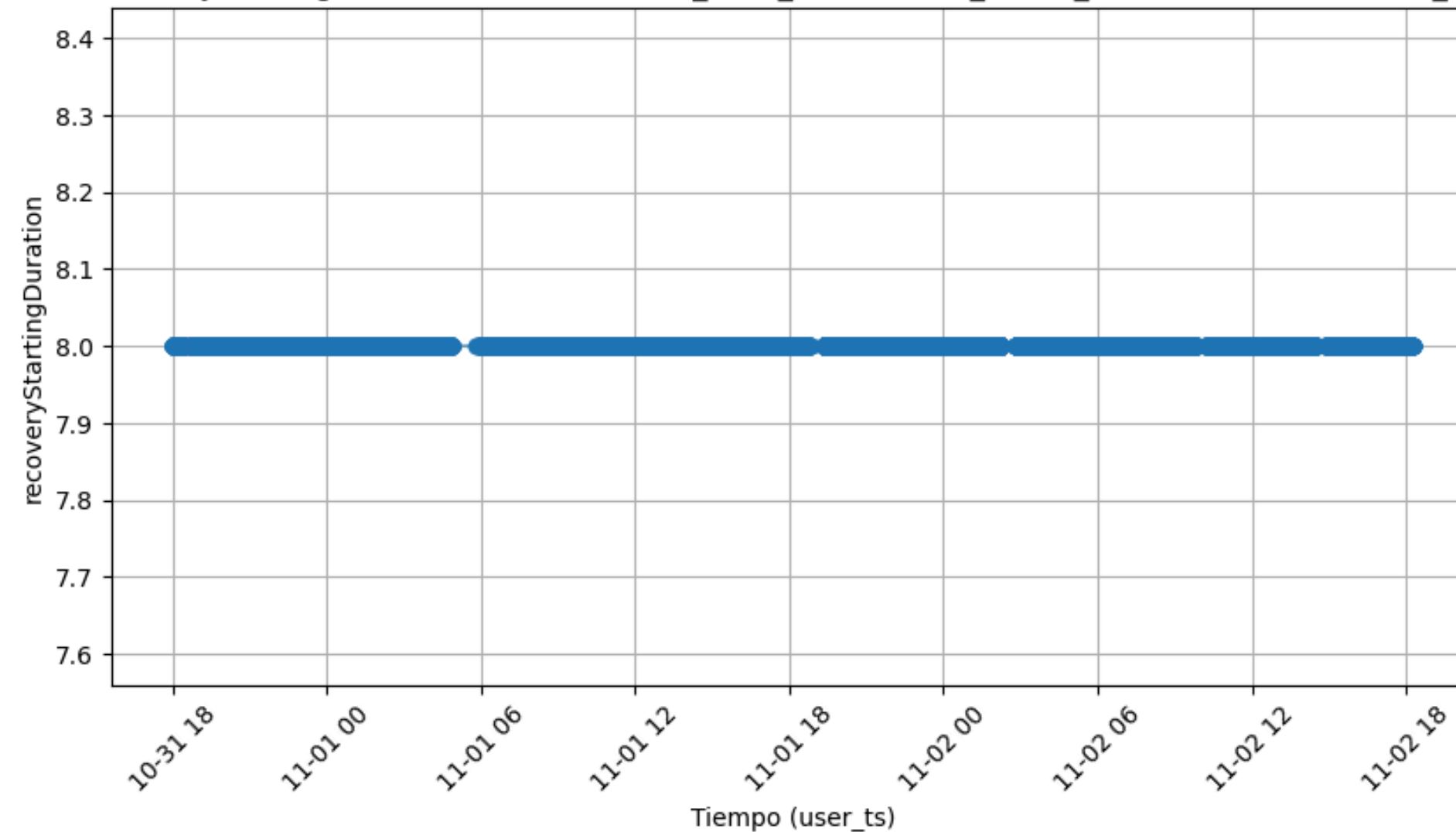
2. Establecimos funciones específicas para un segundo filtrado de las variables, basada en métricas clave:

- Desviación estándar:
- IQR (Rango Intercuartil)
- Número de Outliers
- Proporción de Datos

SELECCIÓN Y LIMPIEZA DE VARIABLES

Ejemplo de variable con desviación estándar muy baja (casi nula) y un IQR de 0.0

Evolución de recoveryStartingDuration en CONTIFORM_MMA_CONTIFORM_MMA1_AirWizardPlusController_0 (primer chunk)



VARIABLES UTILIZADAS

- CONTIFORM_MMA_CONTIFORM_MMA1_Preform TemperatureLayer_1
- CONTIFORM_MMA_CONTIFORM_MMA1_Preform TemperatureLayer_3
- CONTIFORM_MMA_CONTIFORM_MMA1_Preform TemperatureLayer_5
- CONTIFORM_MMA_CONTIFORM_MMA1_Preform TemperatureLayer_7
- CONTIFORM_MMA_CONTIFORM_MMA1_Preform TemperatureLayer_9
- CONTIFORM_MMA_CONTIFORM_MMA1_Final BlowingPressureActualValue_0
- CONTIFORM_MMA_CONTIFORM_MMA1_PressureCompensationChamberPressureActualValue_0
- CONTIFORM_MMA.CONTIFORM_MMA1.WS_Tot _Rej.0

Análisis temporal

Series de tiempo para estudiar fluctuaciones en:
Temperatura.

Presión de soplado.

Presión de la cámara de compensación.

Correlaciones

Matriz de correlación y gráficos de dispersión para identificar relaciones entre variables.

The logo consists of the letters 'EDA' in a bold, black, sans-serif font, centered within a light pink rounded rectangle.

Medidas Estadísticas Iniciales

Media, mediana y desviación estándar para evaluar tendencias.
Mínimos, máximos y percentiles para identificar posibles errores de medición.

Distribución de datos y outliers

Histogramas y boxplots para visualizar patrones y detectar valores extremos.
Uso del rango intercuartil (IQR) y Z-score para identificar outliers.

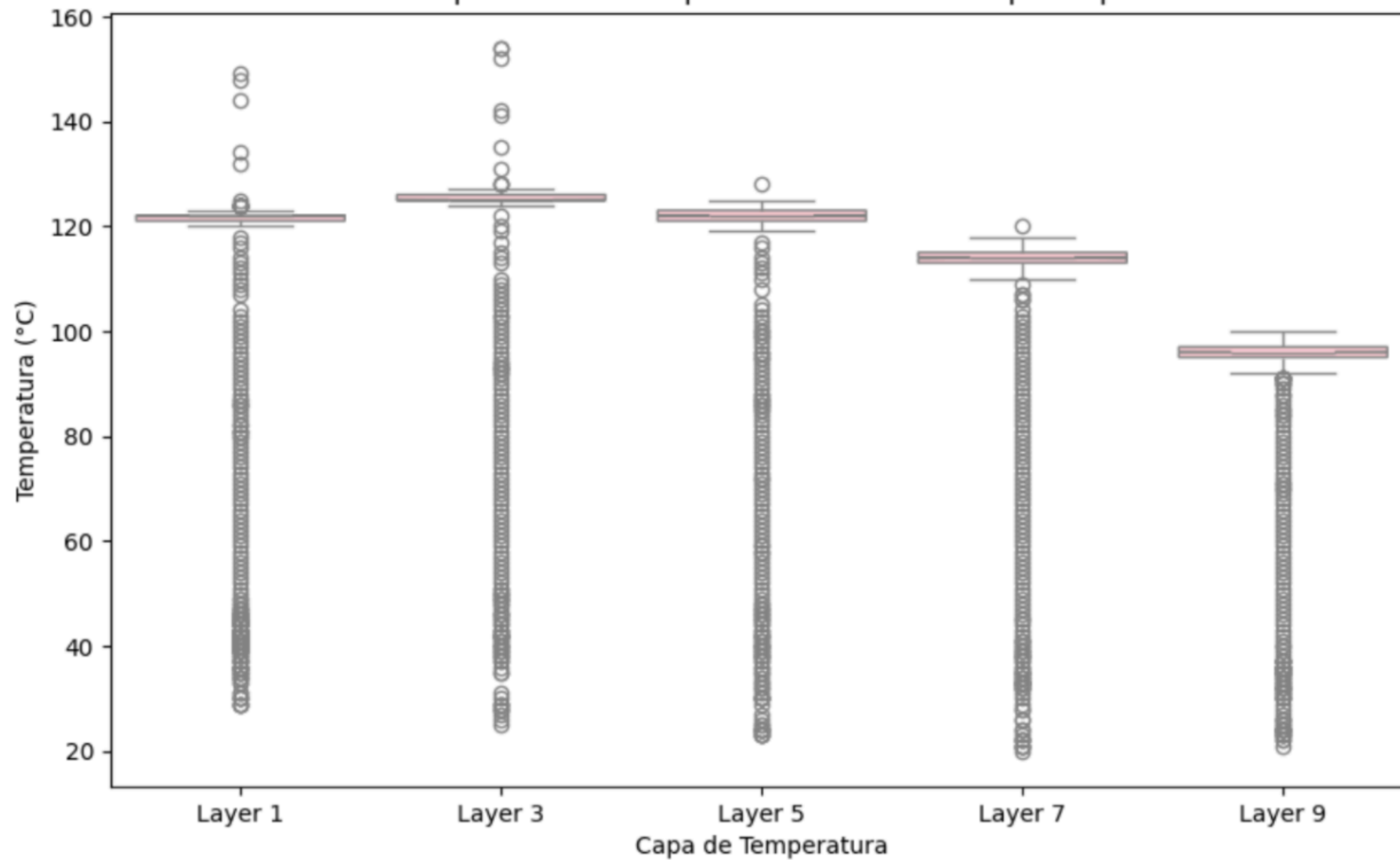
Detección de anomalías

DBSCAN, Isolation Forest, Local Outlier Factor

HIPÓTESIS 1:

“El aumento de la temperatura de la preforma genera un mayor numero de botellas rechazadas”

Comparación de Temperaturas de Preforma por Capa



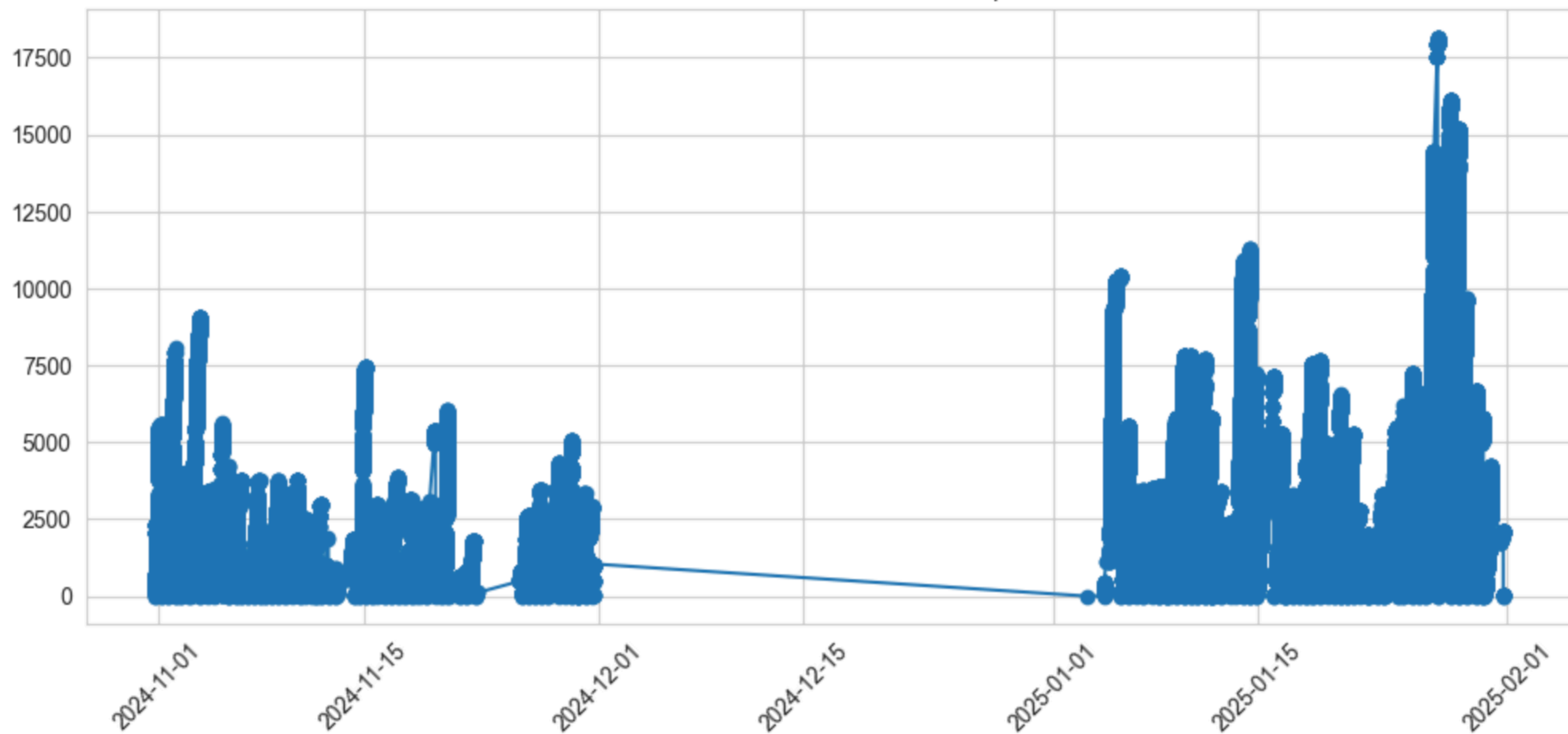
ANÁLISIS DE BOXPLOTS

- Se compararon distribuciones de temperatura en lotes con alto y bajo rechazo.
- Los lotes con alto rechazo presentan mayor dispersión térmica.
- Se identificaron outliers en temperaturas extremadamente bajas ($\sim 20-50^{\circ}\text{C}$).

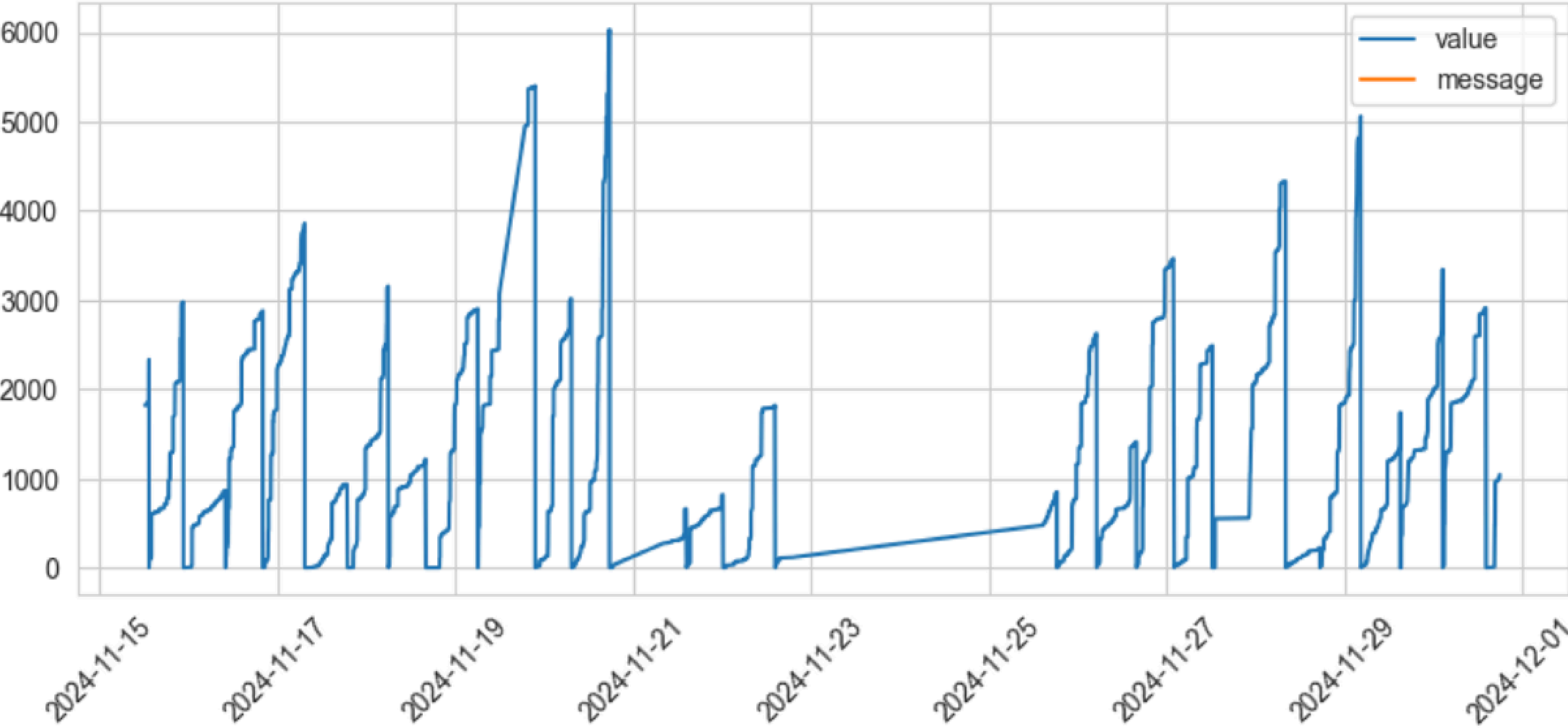
Serie de Tiempo de Temperaturas de Preforma y Total Rechazos (1 Nov - 2 Ene)



Evolución de 'value' en el tiempo



Comportamiento en el intervalo 2024-11-15 12:20:00 - 2024-12-01 12:45:00



SERIES DE TIEMPO

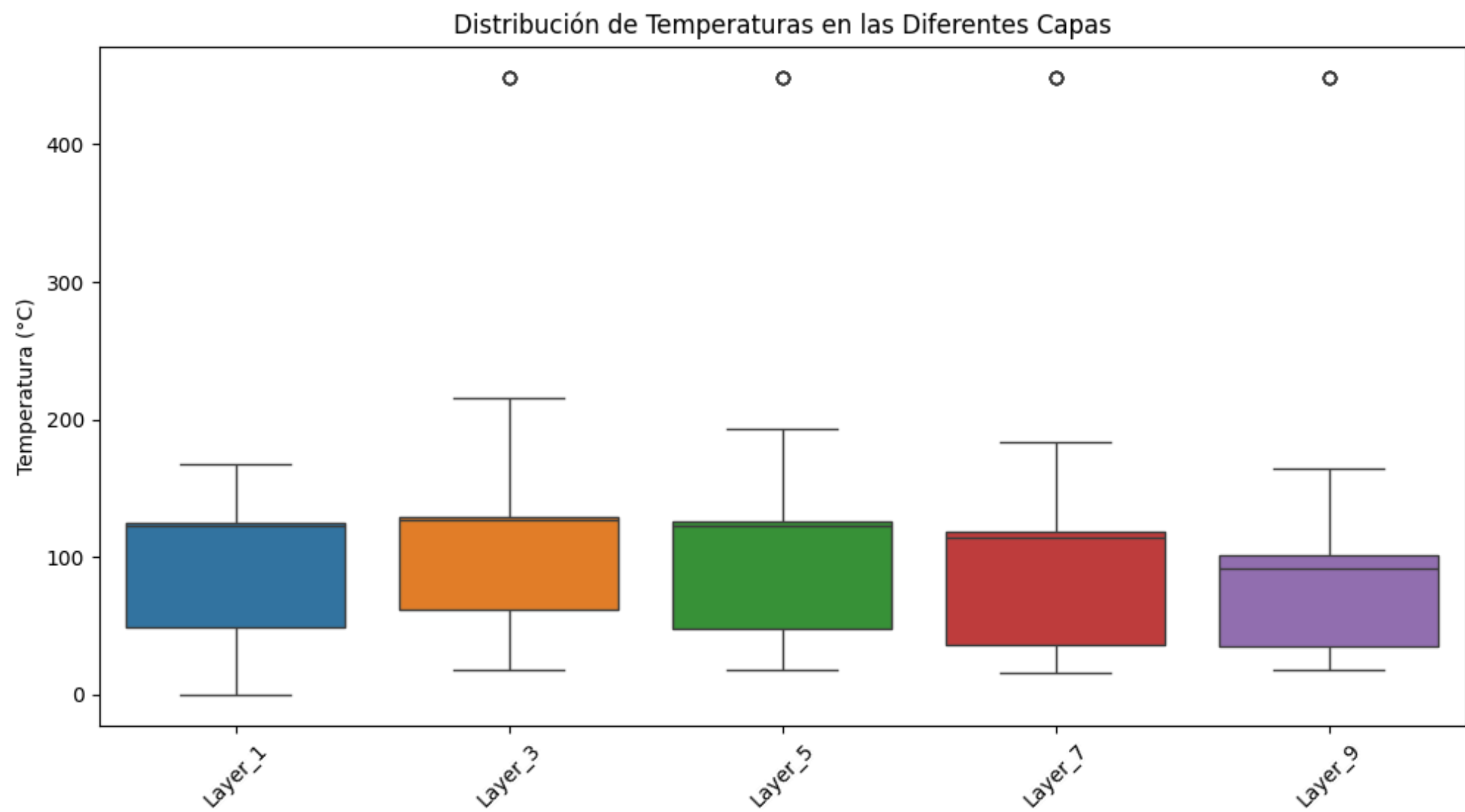
- Se identificó un pico extremo de temperatura el 19 de noviembre de 2024.
- Ese mismo día se registró la mayor cantidad de botellas rechazadas.
- Entre el 23 y 25 de noviembre, la máquina parece haber estado apagada.

CONCLUSIONES

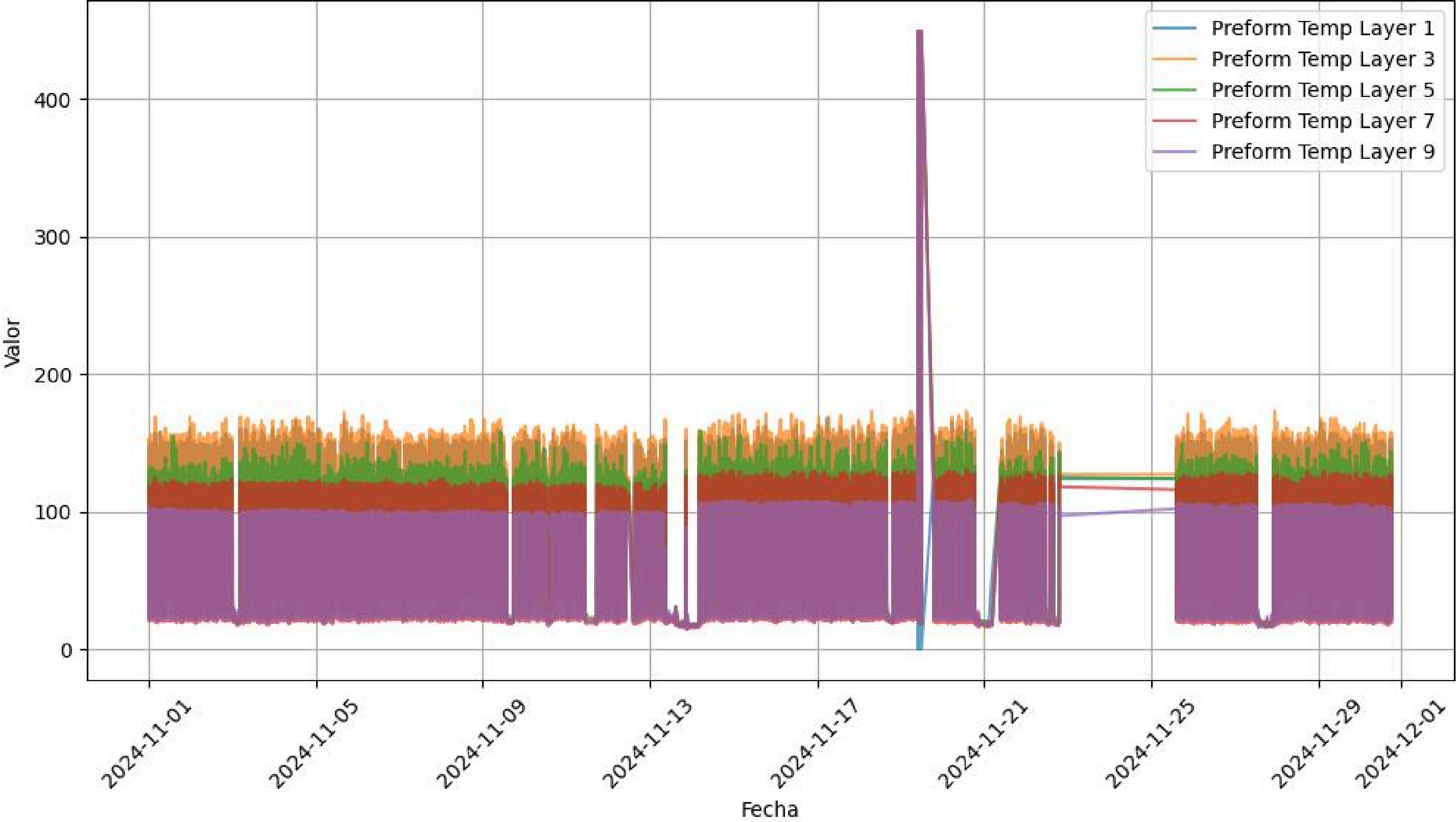
- Fluctuaciones extremas en temperatura pueden estar generando defectos en la preforma.
- El análisis de series de tiempo y boxplots sugiere una correlación entre temperatura y rechazo.

HIPÓTESIS 2

“El desgaste de la sopladora se acelera cuando las temperaturas de la preforma estan fuera de cierto rango óptimo”



Serie de Tiempo de Temperaturas de Preforma y Total Rechazos (1 Nov - 2 Ene)



- **Mínimos:** Algunas capas tienen valores de 0°C, 16°C o 18°C.
- **Máximos:** En todas las capas, hay temperaturas de hasta 449°C.
- Los **valores medios** están entre 70°C y 103°C.
- Layer 9 tiene la **menor temperatura promedio** (70°C), mientras que Layer 3 tiene la **más alta** (103°C).
- Si la diferencia entre capas es muy grande, el proceso de soplado podría ser inconsistente y afectar la vida útil de la máquina.

RELACIÓN CON LA HIPÓTESIS

Se confirma que hay **variabilidad térmica significativa** entre los Layers.

Existen valores extremos, lo que puede afectar la estabilidad del proceso de soplado.

Diferencias de temperatura entre capas pueden causar estrés en la máquina y afectar la calidad del soplado.

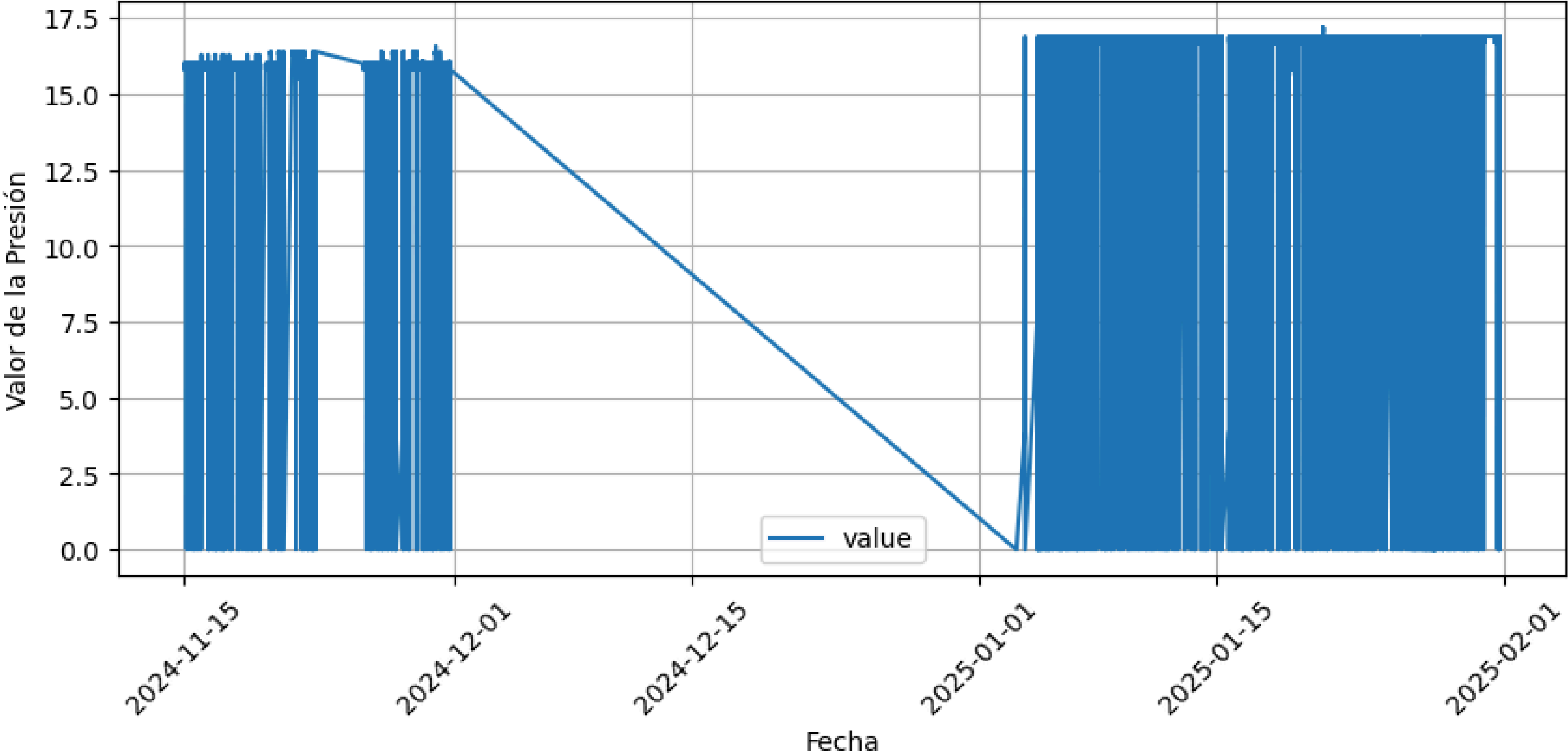
CONCLUSIONES

- Las temperaturas de la preforma no son estables y tienen valores extremos en ciertos momentos --> podría estar causando un mayor desgaste en la sopladora
- Siguiente paso: Cruzar estos datos con registros de mantenimiento, analizar la variabilidad térmica en el tiempo y comparar estos datos con la cantidad de botellas rechazadas.

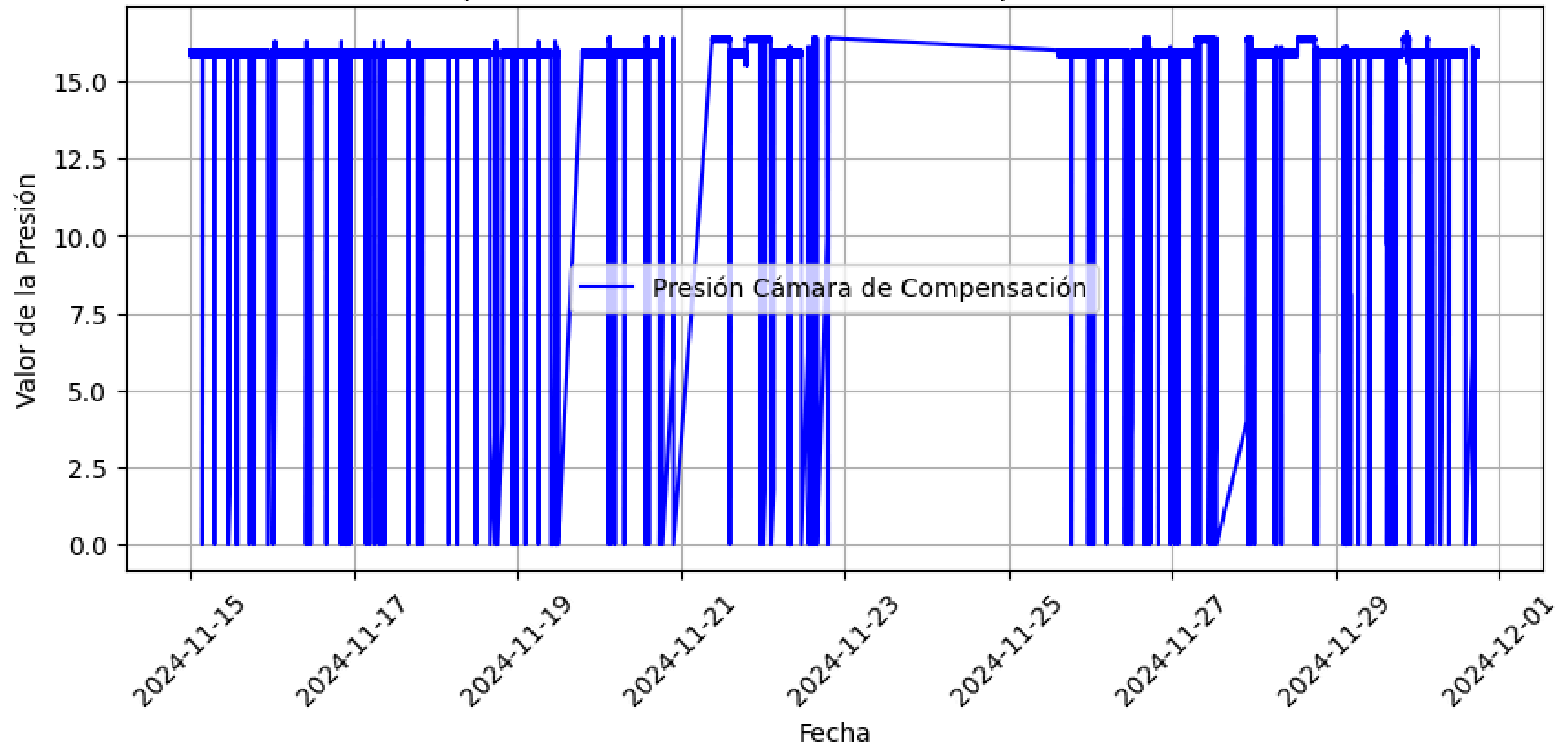
HIPÓTESIS 3

“Variaciones abruptas en la presión de la cámara de compensación aceleran el desgaste de la sopladora y proceso de rechazo de botellas”

Serie de tiempo de Presión de la cámara de compensación (15 Nov - 2 Ene)

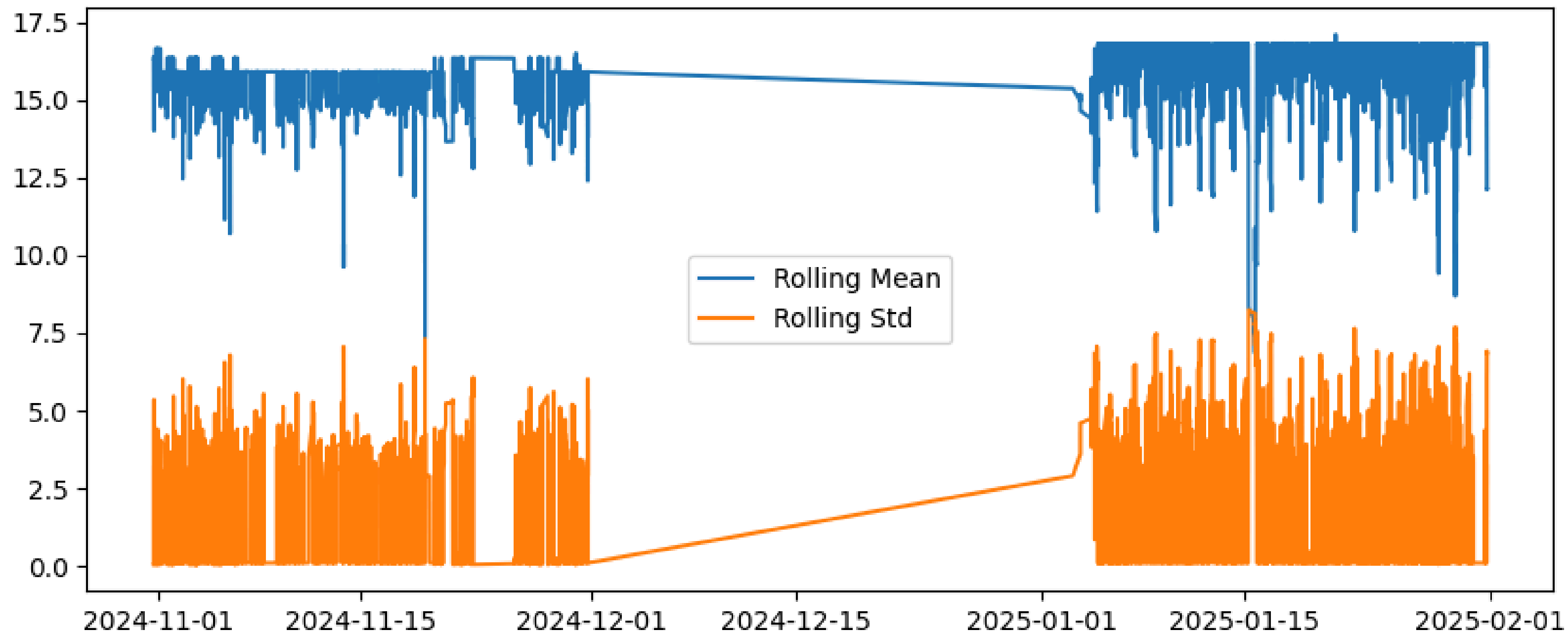


Serie de tiempo de Presión de la cámara de compensación (15 Nov - 2 Ene)

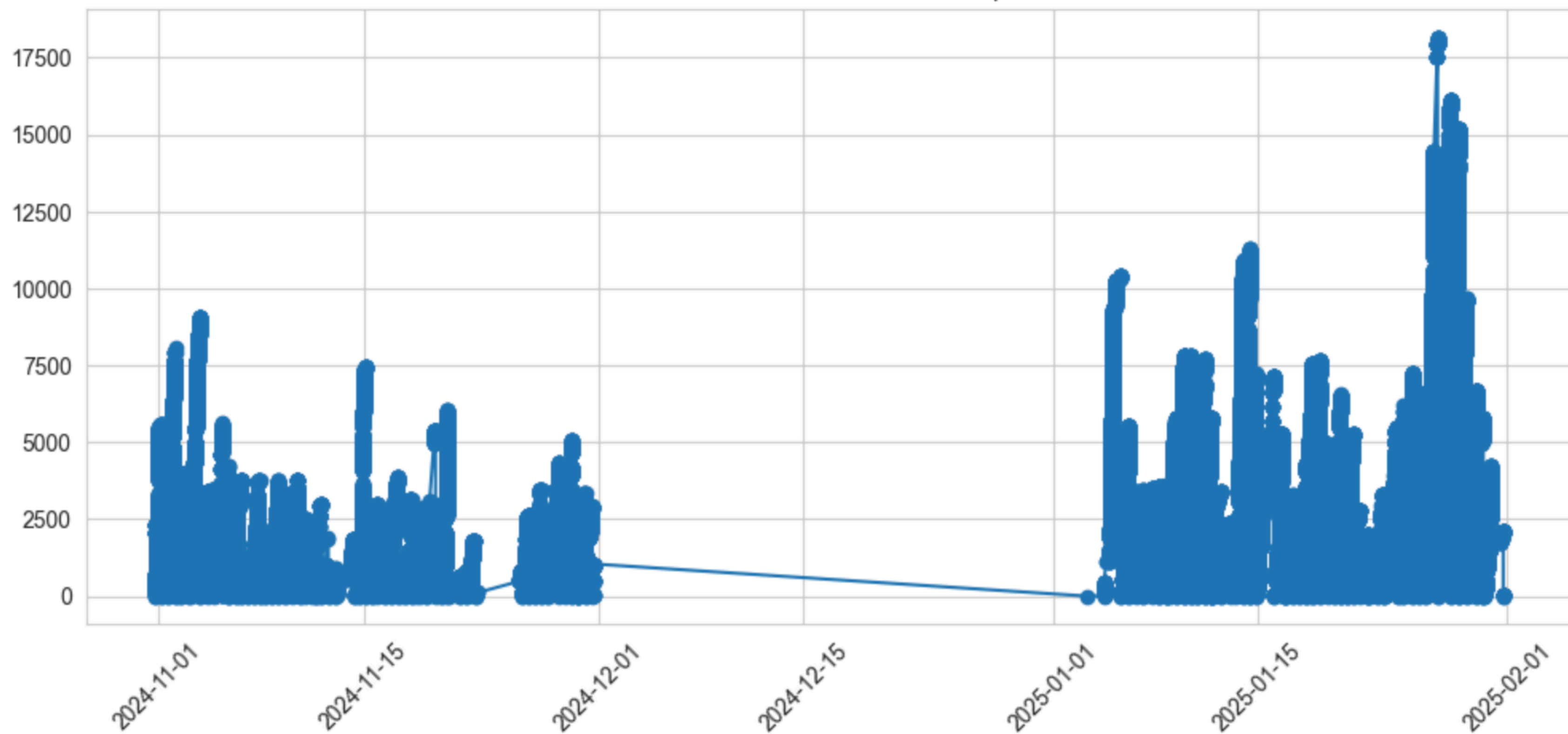


- La presión muestra altas fluctuaciones antes de diciembre y caídas abruptas a cero.
- Al rededor del 23 de noviembre hubo un congelamiento de datos por casi 4 días.
- En diciembre, hay una desconexión seguido de un reinicio en enero.

Rolling Mean & Std de value FinalBlowingPressureActualValue



Evolución de 'value' en el tiempo



- La presión de soplado final muestra fluctuaciones constantes dentro de un rango definido.
- La desviación estándar es alta, lo que indica cambios frecuentes en la presión.
- Cuando crece la media móvil y desviación estándar móvil, también lo hace el número de botellas rechazadas.

CONCLUSIONES GENERALES

- La variabilidad térmica en la preforma está relacionada con un mayor número de botellas rechazadas.
 - Las fluctuaciones extremas y los valores fuera del rango óptimo afectan la calidad del inflado.
- Las temperaturas inestables y los valores extremos podrían estar acelerando el desgaste de la sopladora.
 - Comparar estos datos con registros de mantenimiento.
- Las variaciones abruptas en la presión de la cámara de compensación podrían estar contribuyendo a fallos operativos.
 - Interrupción de datos del 23 de noviembre y la desconexión en diciembre.