



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA DE MINAS

**APLICACIÓN Y EVALUACIÓN DE MODELOS DE MACHINE LEARNING Y
DEEP LEARNING EN LA PREDICCIÓN DE CONSUMO ENERGÉTICO EN
MOLIENDA SAG**

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERA CIVIL DE MINAS

FERNANDA CAROLINA CANALES SOTO

PROFESOR GUÍA:

SEBASTIÁN AVALOS SOTOMAYOR

MIEMBROS DE LA COMISIÓN:

WILLY KRACHT GAJARDO

JULIÁN ORTIZ CABRERA

SANTIAGO DE CHILE

2024

RESUMEN

La industria minera enfrenta grandes desafíos en términos de sostenibilidad y eficiencia energética. La molienda SAG, uno de los procesos que más energía consume, es un área clave para optimizar costos operativos y reducir la huella ambiental. A medida que las operaciones mineras crecen en complejidad y la demanda por eficiencia se intensifica, la capacidad de predecir el consumo energético en tiempo real podría ser una herramienta crucial para ajustar parámetros operativos, permitiendo una mayor eficiencia energética y un control más efectivo de los procesos.

El estudio evalúa la aplicación de modelos de Machine Learning (ML) y Deep Learning (DL) para predecir el consumo energético en un molino SAG, utilizando la metodología propuesta en Avalos et al., (2020), sobre datos de una planta de molienda SAG con características de circuito distintas al estudio de la referencia. El objetivo principal fue aplicar dicha metodología y determinar si la inclusión de nuevas variables, como el estado de los revestimientos y la granulometría, mejora la precisión de las predicciones en este nuevo contexto.

Se implementaron y compararon modelos, como regresión polinómica, máquinas de soporte vectorial (SVR), k-vecinos más cercanos (kNN), redes neuronales multicapa (MLP) y redes de memoria a largo plazo (LSTM). Las variables principales incluyeron el tonelaje procesado, la granulometría del mineral y el desgaste de los revestimientos, de acuerdo con la literatura revisada. Adicionalmente, se probaron variables como la presión en los descansos, evaluando su impacto en la precisión de las predicciones.

Los resultados mostraron que el modelo MLP presentó el mejor ajuste, con un error relativo mínimo cuadrático (rRMSE) de 9.22%, seguido por el modelo LSTM, con un rRMSE de 9.45%. Aunque la adición de variables adicionales ofreció mejoras marginales en algunos casos, no siempre fue significativa en comparación con las variables recomendadas en la bibliografía. Además, se observó que una mayor desviación estándar en los datos se asociaba con errores más altos, y que las diferencias entre los valores característicos de las bases de entrenamiento y validación podían influir en el rRMSE y la correlación de los resultados.

El análisis sugiere que las variables principales son suficientes para obtener predicciones confiables sin agregar complejidad al modelo. La evaluación del estado de los revestimientos fue relevante, aunque no siempre mejoró significativamente el rendimiento. Además, la metodología resultó exportable y efectiva en distintos conjuntos de datos, lo que indica su viabilidad en diferentes contextos operacionales.

Para estudios futuros, se recomienda reducir el error rRMSE y mejorar la correlación entre las predicciones y los datos reales mediante el uso de datos más representativos y diversos. Asimismo, combinar modelos de Machine Learning con modelos físicos tradicionales podría fortalecer la robustez de las predicciones. Evaluar la implementación de estos modelos en sistemas de control en tiempo real permitiría ajustes automáticos para optimizar el consumo energético y mejorar la estabilidad operativa.

TABLA DE CONTENIDO

1. Introducción.....	1
1.1 Objetivos.....	2
1.2 Alcances.....	2
1.3 Estructura de capítulos.....	3
2. Antecedentes.....	4
2.1 Modelamiento en molienda.....	4
2.2 Métodos predictivos en inteligencia artificial.....	5
2.3 Aplicación de modelos IA en la industria minera.....	10
3. Metodología.....	13
3.1 Uso de Python y librerías.....	13
3.2 Entendimiento de los modelos a utilizar.....	13
3.3 Generación de base de datos.....	13
3.4 Modelamiento.....	14
3.5 Evaluación de modelos y comparación de resultados.....	16
3.6 Diagrama de flujo.....	17
4. Aplicación de modelos.....	18
4.1 Caso de estudio.....	18
4.2 Generación de base de datos.....	19
4.3 Modelamiento.....	34
5. Discusión.....	52
6. Conclusiones.....	58
7. Bibliografía.....	60
Anexo.....	62

ÍNDICE DE TABLAS

Tabla 1. Características de equipos base de datos utilizada	18
Tabla 2a. Estadísticas básicas de base de datos pre-limpieza.....	20
Tabla 3. Condiciones limpieza de datos	23
Tabla 4. Estadísticas básicas de base de datos post-limpieza parte 1	23
Tabla 5. Estadísticas básicas de base de datos post-limpieza parte 2.....	23
Tabla 6. Estadísticas básicas de acuerdo a dominio	27
Tabla 7. Est. básicas de acuerdo al uso de bombas para el funcionamiento de 2 molinos de bolas	28
Tabla 8. Campañas de revestimiento.....	30
Tabla 9. Estadísticas básicas por campaña de revestimientos	30
Tabla 10. Definición set de entrenamiento y validación	32
Tabla 11. Promedio consumo energético según bases de entrenamiento.....	34
Tabla 12. Desviación estándar consumo energético según bases de entrenamiento	35
Tabla 13. Promedio variables operacionales según bases de entrenamiento	35
Tabla 14. Desviación estándar variables operacionales según bases de entrenamiento.....	35
Tabla 15. Estadística descriptiva para normalización bases set 1 y set 8.....	47
Tabla 16. Mejores resultados modelamiento base caso 1.....	53
Tabla 17. Mejores resultados modelamiento con variables adicionales caso 2.....	54
Tabla 18. Comparación error mejor caso 1 vs mejor caso 2 por modelo	55
Tabla 19. Comparación modelos machine learning bibliografía caso 3.....	55
Tabla 20. Comparación modelos deep learning bibliografía caso 3.....	56

ÍNDICE DE FIGURAS

Figura 1. Cronología de la inteligencia artificial, aprendizaje de máquina y aprendizaje profundo	5
Figura 2. Ejemplos de parámetro C de valor alto (izquierda) y valor bajo (derecha).	7
Figura 3. Red neuronal del tipo perceptrón simple	8
Figura 4. Estructura célula LSTM	9
Figura 5. Diagrama de flujo metodología	17
Figura 6. diagrama circuito molienda SAG SABC-A	18
Figura 7. Variables principales en el tiempo para base de datos pre-limpieza parte 1	21
Figura 8. Variables principales en el tiempo para base de datos pre-limpieza parte 2	22
Figura 9. Matriz de correlación para base de datos post-limpieza	24
Figura 10. Estimación de densidad: Consumo energético vs tonelaje alimentado SAG	25
Figura 11. Dominios de potencia según número molinos de bolas utilizados	26
Figura 12. Dominios de potencia según número de bombas utilizados	26
Figura 13. Estimación de densidad bivariable para base de datos dominio 2MB	29
Figura 14. Promedio versus desviación estándar según set de datos	31
Figura 15. Consumo de potencia en el tiempo según set de datos	32
Figura 16. Modelamiento caso 1 Regresión Polinomial	37
Figura 17. Modelamiento caso 2 Regresión Polinomial	38
Figura 18. Modelamiento caso 3 Regresión Polinomial	38
Figura 19. Modelamiento caso 1 Vecindad K cercano	39
Figura 20. Modelamiento caso 2 Vecindad K cercano	39
Figura 21. Modelamiento caso 3 Vecindad K cercano	40
Figura 22. Modelamiento caso 1 Regresión con vectores de soporte	41
Figura 23. Modelamiento caso 2 Regresión con vectores de soporte	41
Figura 24. Modelamiento caso 3 Regresión con vectores de soporte	42
Figura 25. Modelamiento caso 1 Perceptrón multicapa	43
Figura 26. Modelamiento caso 2 Perceptrón multicapa	43
Figura 27. Modelamiento caso 3 Perceptrón multicapa	44
Figura 28. Modelamiento caso 1 Memoria corto-largo plazo	45
Figura 29. Modelamiento caso 2 Memoria corto-largo plazo	45
Figura 30. Modelamiento caso 3 Memoria corto-largo plazo	46
Figura 31. Gráfico dispersión modelo PR n=1 Set1vs8	47
Figura 32. Gráfico dispersión modelo kNNR K=1 Set1vs8	48
Figura 33. Gráfico dispersión modelo SVR C=1 Set1vs8	48
Figura 34. Gráfico dispersión modelo MLP n=4 Set1vs8	49
Figura 35. Gráfico dispersión modelo LSTM n=4 Set1vs8	50
Figura 36. Desnormalización set 8.	51
Figura 37. Variación rRMSE según estadísticas básicas	52

ÍNDICE DE ACRÓNIMOS Y ABREVIATURAS

CC: coeficiente de correlación.

DL: Aprendizaje profundo o Deep Learning.

Etto: Base de datos de entrenamiento, considerando variables de entrada para el entrenamiento de acuerdo a metodología planteada (tonelaje alimentado, velocidad de giro, presión de los descansos, variación en tonelaje alimentado y variación en velocidad de giro).

Etto+FTR: Base de datos entrenamiento, considerando variables de entrada Etto más variable de desgaste de revestimiento, correspondiente a tratamiento acumulado desde instalación de cilindro nuevo.

Etto+PS: Base de datos entrenamiento, considerando variables de entrada Etto más variable de granulometría alimentada, correspondiente a porcentaje menor a 2”.

Etto+FTR+PS: Base de datos entrenamiento, considerando variables de entrada Etto más variable de desgaste de revestimiento y variable de granulometría.

EttoReferencia: Base de datos de entrenamiento utilizada en estudio realizado por Ávalos et al., (2020) correspondiente a SAG2.

IA: Inteligencia artificial.

kNNR: Vecindad de k-vecinos más cercanos o K-Nearest Neighbors Regression.

LSTM: Memoria corto-largo plazo o Long-Short Term Memory.

ML: Aprendizaje de máquina o Machine Learning.

MLP: Perceptrón multicapa o Multilayer Perceptron.

PR: Regresión polinomial o Polynomial Regression.

rRMSE: error relativo a la raíz del error cuadrático medio (relative Root Mean Squared Error).

Set i: Base de datos i, siendo un subconjunto i de la base de datos completa que contempla un período de mantenimiento con cambio de revestimientos.

SVR: Regresión con vectores de soporte o Support Vector Machine.

1. INTRODUCCIÓN

El uso de modelos predictivos basados en inteligencia artificial ha avanzado rápidamente en procesos que manejan grandes volúmenes de datos. Estos modelos utilizan el Big Data disponible para crear y reconocer patrones de comportamiento mediante algoritmos computacionales. Esto, junto al aumento de las capacidades de procesamiento computacional, ha hecho posible desarrollar modelos cada vez más complejos, pasando del Aprendizaje de máquinas (Machine Learning) mediante el uso de algoritmos matemáticos de regresión hasta el Aprendizaje profundo (Deep Learning) mediante el uso de redes neuronales.

El campo de las geociencias no ha estado exento del uso de estos modelos, empleando la predicción basada en aprendizaje de máquina para predecir y describir sistemas complejos mediante el análisis de patrones multivariantes.

Dentro del campo de las geociencias, específicamente en la minería, se encuentra el proceso de conminución de minerales, donde a partir de las características propias de cada mineral como composición, granulometría, dureza, entre otros, se obtiene un consumo energético requerido para su disminución de tamaño.

El proceso de conminución de minerales corresponde a uno de los procesos industriales con el uso de energía más intensivo dentro del proceso minero, donde en particular la molienda es conocida por su baja eficiencia energética. Es por esto que es crucial contar con sistemas de predicción que, basados en los valores observados en el molino, puedan estimar el consumo energético en un intervalo de tiempo acotado, lo suficiente para ser incorporado en sistemas de control que automatizan la toma de decisiones para mejorar la eficiencia energética y prevenir sobrecargas del molino.

Desde los años 60, con la creación de modelos empíricos y teóricos de consumo energético basados en las dimensiones, distribución de la carga y tonelaje de alimentación de los molinos, como los de Bond et al. (1961) [2] y Austin et al. (1972) [4], hasta la inclusión de variables geometalúrgicas, se ha avanzado significativamente en el desarrollo de modelos predictivos. No obstante, aún existe margen para mejorar su grado de predicción, con el objetivo de optimizar su aplicación en el control de procesos a tiempo real.

Es por esto, y dada la gran cantidad de datos que la industria posee hoy en día, que se ha intensificado la investigación en el uso de algoritmos de Machine Learning y Deep Learning para el modelamiento y predicción a corto plazo de diversas variables operacionales. Esta evolución en el análisis de datos no solo ha fortalecido los sistemas de control, sino que también ha proporcionado indicadores clave para la toma de decisiones operacionales. Este enfoque avanzado resulta en sistemas más robustos y eficientes, mejorando significativamente la gestión y optimización de procesos en entornos industriales de alta complejidad debido a la cantidad de variables operacionales a considerar.

El presente trabajo considera un conjunto de métodos predictivos entrenados para inferir el consumo energético en molienda SAG. Diferenciándose de trabajos anteriores [1], no solo se examina la precisión de las predicciones en este nuevo contexto, sino que también aborda la posibilidad de incluir variables operativas lo que se espera enriquezca y potencialmente mejore la

precisión de las predicciones. Al hacerlo, este trabajo no solo se alinea con el estado del arte en la predicción del consumo energético en molinos SAG, sino que también busca expandirlo, proporcionando nuevos conocimientos y posibles rutas para optimizaciones futuras en la eficiencia energética y operativa de estos procesos críticos en la industria minera.

1.1 OBJETIVOS

El desarrollo del presente trabajo considera los siguientes objetivos y alcances.

1.1.1 OBJETIVO GENERAL

- Entrenar, evaluar y validar el uso de modelos de Machine Learning y Deep Learning en la predicción de consumo energético en molienda SAG

1.1.2 OBJETIVOS ESPECÍFICOS

- Definir set de datos a utilizar en modelos predictivos, indicando discretización temporal, variables a utilizar y definición de estado estacionario.
- Identificar influencia del estado de revestimientos en las predicciones de consumo energético.
- Revisar el desempeño de metodologías de predicción para modelamiento utilizando Machine Learning (modelos PR, kNNR, y SVR) y Deep Learning (modelos MLP y LSTM) en datos operacionales de molienda SAG.
- Comparar desempeño de modelos con respecto a bibliografía para el uso de modelos similares con el propósito de determinar relaciones.
- Definir los mejores modelos y evaluar su desempeño al agregar nuevas variables.

1.2 ALCANCES

- Los datos utilizados se ubican en horizonte temporal entre 2016 y 2020, incluyendo datos operacionales de molino SAG y datos de operatividad de equipos aguas abajo con datos obtenidos cada 5 minutos.
- Se utilizarán modelos de predicción pertenecientes a estudio previo [1], pero con separación de bases únicamente de entrenamiento y validación, sin incluir base de prueba.
- Dada la cantidad de modelos, se decide realizar este caso considerando sólo el soporte temporal de 30 minutos.
- Para modelos de Deep Learning se mantendrán hiperparámetros de bibliografía [1].
- Se utilizará método de mínimos cuadrados y correlación para selección de mejores resultados.

1.3 ESTRUCTURA DE CAPÍTULOS

La estructura del trabajo se encuentra separada en 7 capítulos, descritos a continuación:

Capítulo 1. Introducción: Se presenta la introducción, objetivos y alcances de la memoria de título.

Capítulo 2. Antecedentes: Realiza una recopilación de antecedentes asociados a este trabajo. Para esto se comienza con una revisión bibliográfica la definición de inteligencia artificial y su derivación en modelos de Machine learning y Deep learning. Se describe brevemente los modelos utilizados y el uso de aprendizaje de máquina para las geociencias y la minería. Posteriormente, se presentan antecedentes propios de la industria y la manipulación de variables a tiempo real.

Capítulo 3. Metodología: Presenta la metodología utilizada en el preprocesamiento de la base de datos y predicción de consumo energético basado en variables operacionales. Se describen y explican las variables operacionales, y el impacto que ellas suponen sobre el consumo energético en molienda SAG.

Capítulo 4. Aplicación de modelos: Desarrollo de casos de estudio. Generación de base de datos, el trabajo de la información obtenida y aplicación de la metodología previamente definida en el pronóstico de consumo energético.

Capítulo 5. Discusión: Se discuten los resultados en dos secciones, la primera acerca de los principales resultados obtenidos y la selección de los mejores modelos para cada método utilizado junto con comparar los resultados obtenidos con la bibliografía. En la segunda sección se discute acerca de la aplicación del modelamiento del consumo de potencia a tiempo real en la industria.

Capítulo 6. Conclusiones: Finalmente, se presentan las conclusiones asociadas al trabajo realizado.

Capítulo 7. Bibliografía: Presentación de la bibliografía utilizada en el trabajo.

2. ANTECEDENTES

2.1 MODELAMIENTO EN MOLIENDA

A lo largo de los años, son diversos los modelos teóricos y empíricos que se han estudiado para la predicción del consumo energético y/o energía específica en la etapa de conminución de mineral, especialmente en la molienda.

En el año 1961 se da a conocer uno de los primeros modelos de los que se tiene registro en la historia moderna, correspondiente al modelo de Bond et al. (1961) [2] para la predicción de consumo energético, el cual se basaba en el desempeño de la disminución de tamaño de la partícula de mineral, calculando un índice de trabajo operacional que permitía obtener la cantidad de potencia utilizada. Este modelo aún es utilizado para la evaluación y diseños de circuitos de molinos de bolas hoy en día.

Posteriormente Hogg & Fuerstenau [3] en 1972 proponen una ecuación de potencia basada en la forma de la carga al interior del molino, donde para el riñón generado se define un talón y un hombro en giro, además de un ángulo de reposo de la carga. Austin [4] desarrolla en 1990 un modelo basado en las propuestas de Bond y Hogg&Fuerstenau, donde se encargó de ajustar los parámetros mediante el uso de datos de molinos industriales.

En 1996 Morrell [5] propone una ecuación para el cálculo del consumo específico de energía basado en características del mineral, llenado del molino, velocidad crítica y distribución de la carga al interior del molino, en la cual se considera también el comportamiento de la carga en los extremos del molino, así como las pérdidas de energía mecánicas, basado en el estudio de 75 plantas industriales, para la predicción del consumo energético. Posteriormente estos modelos serían modificados, agregando nuevas variables y fórmulas para mejorar los resultados. En los últimos años se ha trabajado en incluir variables geometalúrgicas a estos modelos para mejorar la predicción, como por ejemplo Morrell en 2009 [6] donde se incluye el tamaño de alimentación F80 en la predicción y A. Casali [7] en 2015 propone un modelo empírico donde utiliza como variable de entrada la granulometría de alimentación a molino SAG mediante la inclusión del porcentaje de partículas intermedias.

En muchos proyectos de ingeniería se utilizan estos métodos, que en general incluyen la utilización de alguna de las ecuaciones planteadas para predicción de consumo energético o energía específica basadas en las dimensiones del molino, el nivel y la densidad de la carga interna, tonelaje de alimentación y velocidad crítica, siendo utilizados de forma exclusiva para el diseño y/o evaluación de los circuitos de conminución, sin conseguir un error bajo al ser utilizados con datos de operación a corto plazo.

Hoy en día estos métodos no son suficientes para ser utilizados en la optimización y automatización de procesos, donde se requiere un alto grado de certeza en la predicción para ser utilizado en control de procesos a tiempo real.

Los avances tecnológicos actuales, aumento de capacidad de almacenamiento y aumento de sensorización de plantas concentradoras permiten obtener y guardar gran cantidad de datos, permitiendo generar bases de datos históricas con el comportamiento de variables operacionales y de respuestas en molinos SAG en intervalos cortos de tiempo.

Es debido a esto que el aprendizaje de máquina aparece como una alternativa práctica para ser utilizado en la generación de modelos de predicción de variables en la industria a mediano y corto plazo, e incluso a tiempo real, ya que permiten a partir de los datos almacenados que las resoluciones temporales pueden modificarse para generar modelos de predicción desde minutos a incluso segundos, y generar modelos de aprendizaje que sean capaces de identificar y reproducir patrones o tendencias entre los datos que no son fácilmente reconocibles con otros métodos, y que es posible capturar a partir de la ciencia de datos.

2.2 MÉTODOS PREDICTIVOS EN INTELIGENCIA ARTIFICIAL

Existen múltiples definiciones para la Inteligencia Artificial (IA), siendo lo más común describirla como la inteligencia humana exhibida por máquinas o método por el cual se permite que las máquinas piensen sin la intervención humana, permitiéndole a la máquina ser capaz de tomar sus propias decisiones.

La disponibilidad de grandes conjuntos de datos, el desarrollo de algoritmos efectivos y eficientes, y el acceso a la computación de alto rendimiento han dado como resultado el éxito del aprendizaje de máquinas en los últimos años.

El aprendizaje de máquina o *Machine Learning* es una subclase de la IA, donde se le agrega la cualidad de aprendizaje a la máquina, para así mejorar la toma de decisión [12]. Suele utilizar algoritmos de aprendizaje estadístico para construir sistemas que puedan mejorar a través de la experiencia sin volver a ser programados.

El aprendizaje profundo o *Deep Learning* corresponde a una subclase del aprendizaje de máquina, donde los métodos de aprendizaje se vuelven más complejos, inspirándose en la forma que tienen las neuronas, organizando y clasificando la información durante el proceso de aprendizaje.

De esta forma, los métodos asociados a la inteligencia artificial se encuentran autocontenidos, tal como es mostrado en Figura 1.

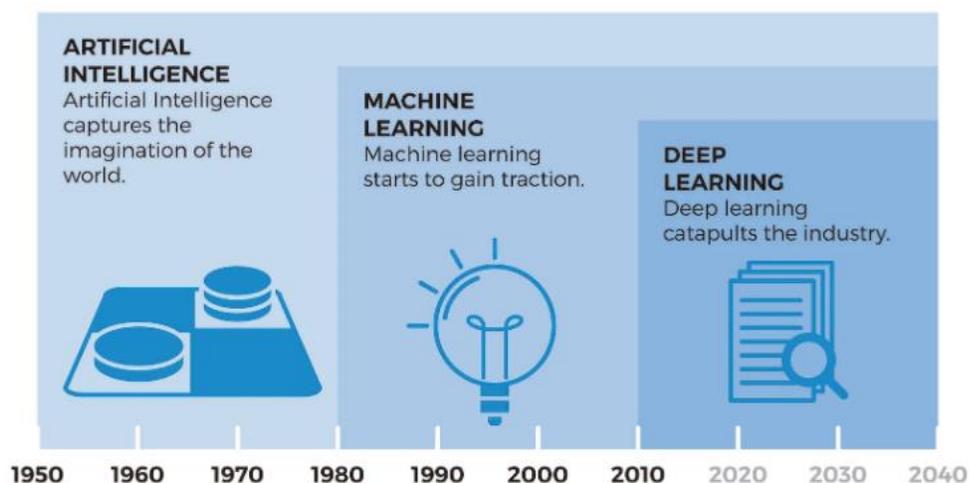


FIGURA 1. CRONOLOGÍA DE LA INTELIGENCIA ARTIFICIAL, APRENDIZAJE DE MÁQUINA Y APRENDIZAJE PROFUNDO

La mayoría de los métodos de aprendizaje profundo utilizan arquitecturas de redes neuronales con varias capas ocultas por lo que se conocen como redes neuronales profundas. Estos métodos son entrenados mediante la iteración en varias capas de procesamiento, creando múltiples niveles de abstracción para la representación de los datos.

Para el modelamiento mediante aprendizaje profundo es importante conocer los conceptos de lote (*batch*) y época (*epoch*), hiperparámetros utilizados para el ajuste de los modelos. El número de lotes es un hiperparámetro que controla la cantidad de muestras del set de entrenamiento para trabajar antes de que se actualicen los parámetros internos del modelo. Por otro lado, el número de épocas es un hiperparámetro que controla el número de pases completos (iteraciones) que se le realiza a la base de entrenamiento completa. Estos valores deben ser obtenidos de forma particular para cada proyecto de modelamiento, para evitar así un sub o sobre entrenamiento de las bases.

Dentro de los modelos de aprendizaje de máquina a utilizar en el estudio se encuentran:

- **Regresión polinomial (PR):**
Modelo de regresión lineal en que se genera un polinomio de n ésimo grado. Para el modelamiento se ajustan las constantes del polinomio, donde sólo se debe seleccionar el grado con el que se trabajará. La regresión polinomial es sensible a los valores *outlier*, por lo que la presencia de estos puede afectar negativamente el rendimiento.
- **Vecindad de los k-vecinos más cercanos (k-NNR):**
Modelo de regresión basado en el algoritmo de clasificación del mismo nombre, el cual determina la clase de un punto mediante diversas ecuaciones que consideran sus pares más cercanos. De esta manera, en la regresión, al nuevo punto se le asigna un valor según su similitud con los puntos del conjunto de entrenamiento, en función de la cantidad de vecinos especificados.
- **Regresión con vectores de soporte (SVR):**
Modelo de regresión cuyo objetivo es la generación de un hiperplano en un espacio de n dimensiones que clasifique los puntos de datos. Los puntos de datos más cercanos a cada lado del hiperplano se denominan vectores de soporte. En esta regresión, el mejor ajuste se logra cuando el hiperplano abarca el mayor número de puntos. A diferencia de la regresión polinomial, este modelo es robusto frente a valores *outliers*.
Para el modelamiento, se debe asignar un parámetro de regularización, conocido típicamente como C , el cual indica qué tanto se quiere evitar clasificar erróneamente un valor. Para valores altos de C , el entrenamiento generará un hiperplano de margen más pequeño con los puntos adyacentes que tendrá un mejor desempeño en la clasificación. Por el contrario, para valores muy pequeños de C , el entrenamiento generará un hiperplano de margen más grande, lo que puede resultar en puntos mal clasificados. Se debe probar con distintos valores de C para generar el modelo con un mejor grado de predicción. La Figura 2 muestra ejemplos de cómo el parámetro C influye en el margen del hiperplano en la regresión con vectores de soporte. A la izquierda se observa un valor alto de C y a la derecha un valor bajo.

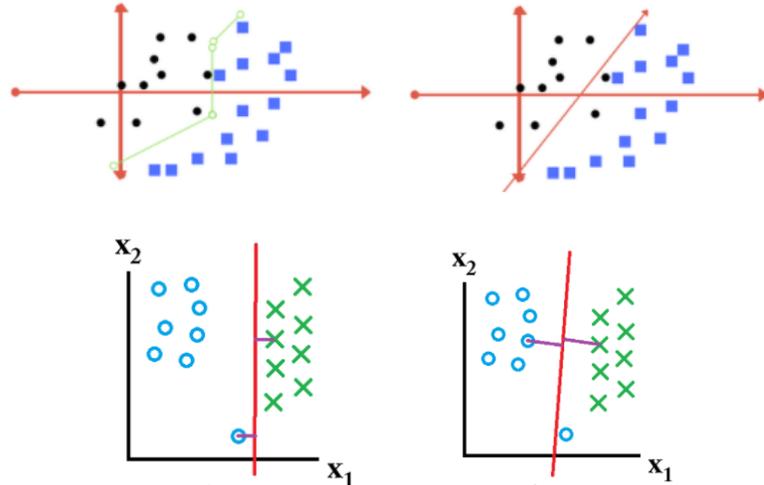


FIGURA 2. EJEMPLOS DE PARÁMETRO C DE VALOR ALTO (IZQUIERDA) Y VALOR BAJO (DERECHA).

Además, existe un segundo hiperparámetro *epsilon* que define la tolerancia al error. Controla el ancho de la banda dentro de la cual las predicciones se consideran precisas y no se penalizan. Su valor puede calcularse directamente a partir de la desviación estándar de los datos y el número de muestras en el conjunto de datos. Un *epsilon* mayor permite un margen de error más amplio, mientras que un *epsilon* menor exige predicciones más cercanas a los valores reales.

Dentro de los modelos de aprendizaje profundo a utilizar en el estudio se encuentran:

- Perceptrón multicapa (MLP)

Un perceptrón multicapa es una clase de red neuronal artificial de retropropagación. Un MLP posee al menos 3 capas, una de entrada, salida y una oculta.

- Capas de Entrada y Salida: La capa de entrada recibe los datos de entrada, mientras que la capa de salida produce el resultado o predicción del modelo.
- Capas Ocultas: Las capas ocultas, que están entre la entrada y la salida, son donde ocurre la mayor parte del procesamiento a través de una combinación lineal seguida de una función de activación no lineal como Rectified Linear Units (ReLU), Sigmoide (Sigmoid) o tangente hiperbólica (TanH).
- Función de Activación: La función de activación en cada neurona determina la salida en base a sus entradas ponderadas. Por ejemplo, la función sigmoide se puede expresar como

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

La propagación hacia atrás o retropropagación es la práctica de ajustar los pesos de una red neuronal en función de la tasa de error obtenida en la época anterior (iteración). Este proceso implica calcular el gradiente de la función de pérdida respecto a cada peso mediante el

algoritmo de descenso de gradiente. El ajuste adecuado de los pesos le permite reducir las tasas de error en las siguientes iteraciones y obtener un modelo más confiable.

El descenso de gradiente se realiza típicamente usando variantes como el Gradiente Descendente Estocástico (SGD), Adam o RMSprop. La ecuación básica del descenso de gradiente es:

$$w_{nuevo} = w_{virjo} - \eta * \nabla J(w)$$

donde η es la tasa de aprendizaje y $\nabla J(w)$ es el gradiente de la función de pérdida.

Para el modelamiento en un modelo de perceptrón multicapa, se debe indicar la cantidad de capas ocultas que tendrá nuestro entrenamiento, los cuales corresponden a sus parámetros críticos. El número de neuronas y la arquitectura en cómo estas están entrelazadas explican la complejidad de la red neuronal. A mayor complejidad de la red, mayor es la posibilidad de que extraiga los patrones subyacentes en la base de datos de entrenamiento. No hay una regla fija para esta elección, y a menudo se determina mediante experimentación y validación cruzada.

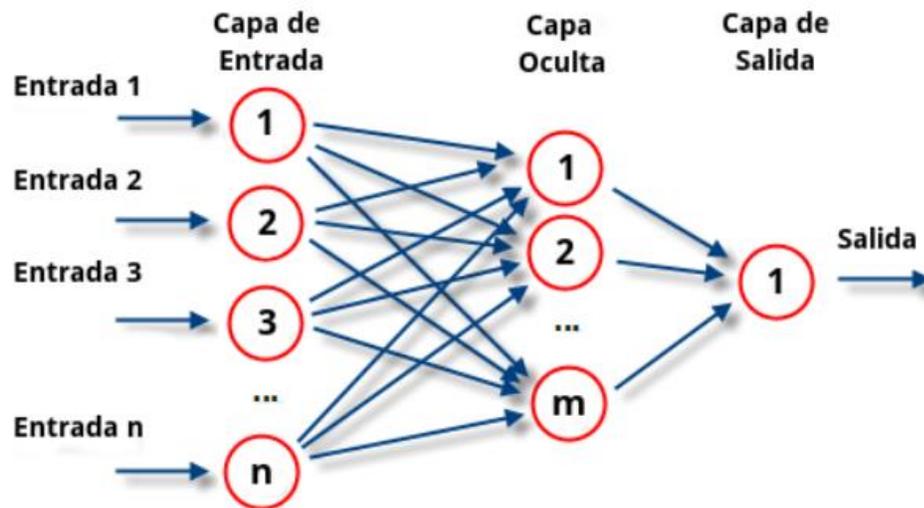


FIGURA 3. RED NEURONAL DEL TIPO PERCEPTRÓN SIMPLE

En Figura 3, se muestra una Red Neuronal Artificial de tipo perceptrón simple con n neuronas de entrada, m neuronas en su capa oculta y una neurona de salida.

- Memoria corto-largo plazo (LSTM)

El modelo de memoria a corto y largo plazo está basado en una arquitectura de red neuronal recurrente (RNN) que permite que la información persista durante el entrenamiento. A diferencia de las RNN tradicionales, los modelos LSTM son capaces de recordar información durante largos períodos de tiempo, lo que los hace idóneos para aplicaciones como el procesamiento de lenguaje natural y series temporales.

Al igual que el perceptrón multicapa, un LSTM tiene al menos tres capas, pero en este modelo las capas ocultas pueden retener la información anterior y utilizarla para procesar la entrada actual.

Los LSTM son ideales para tareas que implican dependencias temporales o secuenciales, como el procesamiento de lenguaje, la predicción de series temporales y el reconocimiento de voz. Su capacidad para manejar secuencias de longitud variable los hace versátiles para el modelamiento de datos temporales en diversos campos.

- Capas y Estructura: Un LSTM típicamente consta de una capa de entrada, varias capas ocultas LSTM y una capa de salida. Cada capa LSTM contiene celdas con una estructura de puertas especializadas que controlan el flujo de información.
- Puertas de la Célula LSTM: Las puertas en cada célula LSTM —puerta de olvido, puerta de entrada y puerta de salida— controlan el mantenimiento y la actualización del estado de la celda, permitiendo que la red retenga o descarte información a lo largo del tiempo.
- Ecuaciones de un LSTM: Las operaciones dentro de una célula LSTM pueden describirse con un conjunto de ecuaciones matemáticas que incluyen operaciones de suma, multiplicación y funciones de activación como la sigmoide y tanh.

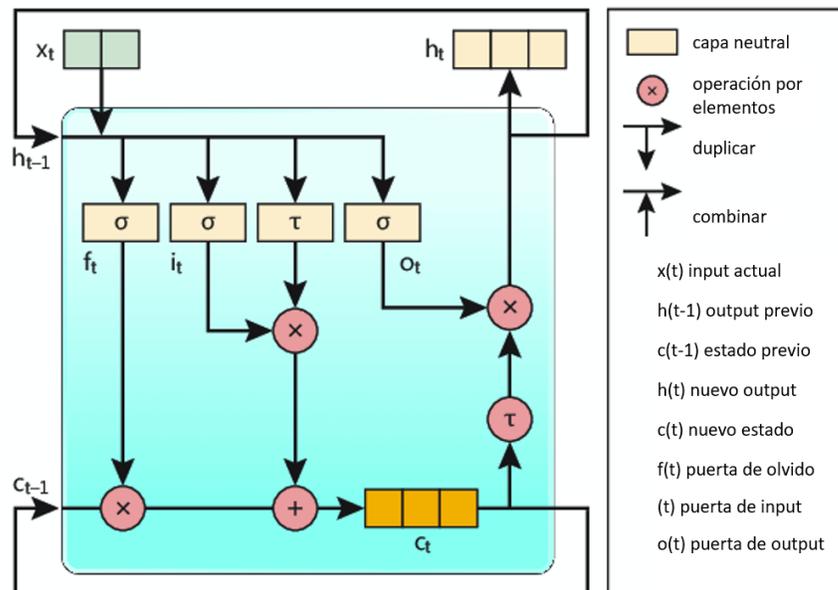


FIGURA 4. ESTRUCTURA CÉLULA LSTM

Al igual que en modelos MLP se debe definir el número de capas que utilizarán las células LSTM y el número de neuronas en cada una. Esta configuración depende en gran medida de la complejidad de la tarea y los datos disponibles. La Figura 4 ilustra la arquitectura de una célula LSTM.

2.3 APLICACIÓN DE MODELOS IA EN LA INDUSTRIA MINERA

En la actualidad modelos basados en Big Data, aprendizaje de máquina y aprendizaje profundo se vuelven valiosas herramientas para inferir nueva información a partir de los datos disponibles en el sistema, donde la minería no es un campo exento de esto. Tanto para los campos de exploración de depósitos minerales, modelamiento a partir de sondajes, optimización de modelos de bloques de extracción y procesamiento, control automático de los procesos, predicción del comportamiento de variables operacionales, entre otras áreas, se ha investigado acerca del uso y aplicación de estos modelamientos que difieren del concepto clásico teórico-empírico que se manejaba.

Geociencias

Dentro del campo geológico-minero, las Geociencias son responsables de diseñar-planificar y ejecutar el proceso de obtención de información de todos los parámetros relevantes que pueden afectar el proceso de toma de decisiones de un negocio minero desde el principio de su conceptualización, hasta la puesta en marcha y operación. Los cambios que han ocurrido en las faenas, proyectos y mercados de la minería a nivel económico, tecnológico y físico han implicado una presión por mejoras en la construcción de modelos geológicos más confiables y acertados, que mejoren la comprensión de los procesos basados en las Geociencias (geoquímica, estimación de recursos, geometalurgia, entre otros).

A pesar de esto, se debe destacar que las aplicaciones de modelos basados en aprendizaje de máquina en las geociencias no se limitan al campo minero, sino a campos relacionados con cambio climático, contaminación, disponibilidad de agua, comida y materias primas, entre otros, los cuales han extensamente investigados a lo largo de los años, donde para cada una de estas materias es importante contar siempre con el apoyo técnico de quién maneje la información de campo, convirtiéndose en un espacio colaborativo entre la ciencia de datos y especialistas de cada área [8][9].

El uso de ML en geociencias está transformando la manera en que los científicos entienden los procesos terrestres y mejorando la exploración y explotación de recursos minerales.

Geometalurgia

La IA se puede utilizar para predecir propiedades clave del mineral, como la dureza, la respuesta a la molienda y la lixiviación, basándose en datos geológicos y de ensayos metalúrgicos. Los modelos geometalúrgicos basados en IA facilitan la toma de decisiones informadas en todas las etapas de la minería, desde la exploración hasta la producción.

El trabajo de Lishchuk, Lund y Ghorbani [10] investiga cómo diferentes métodos de machine learning pueden ser utilizados para integrar datos de proceso dispersos en un modelo espacial geometalúrgico. Este enfoque mejora significativamente la capacidad de predecir la respuesta del proceso mineral basándose en la ubicación de la muestra en un yacimiento, lo que resulta crucial para la planificación eficiente de la producción minera.

Por otro lado, Moraga, Kracht y Ortiz [11] presentan un estudio sobre la modelación de procesos para determinar la mezcla y la distribución del tiempo de residencia en plantas de procesamiento mineral. Esta investigación destaca la importancia del modelamiento geometalúrgico y dinámica

en la evaluación del rendimiento de las plantas, utilizando la distribución del tiempo de residencia para anticipar los efectos de las perturbaciones en la alimentación y comprender cómo las variaciones afectan la composición final del producto.

Ambos estudios resaltan el impacto transformador de la IA y el machine learning en la geometalurgia, subrayando cómo estas tecnologías avanzadas pueden mejorar la eficiencia y la sostenibilidad en el sector minero, además de poder anticiparse a los resultados a obtener.

Minería

En el sector minero, el desarrollo y la aplicación de la inteligencia artificial (IA) se han enfocado en la automatización de operaciones, la optimización de la planificación estratégica y la mejora del modelamiento de leyes. La integración de la IA en las operaciones mineras ha generado innovaciones significativas que han mejorado tanto la productividad como la eficiencia energética. Ramezani y Tafazoli (2021) [13] destacan cómo la implementación de sistemas de monitoreo y control en tiempo real facilita la recopilación y análisis de datos precisos, mejorando así la toma de decisiones operativas.

Los avances en la planificación minera son igualmente notables, como se expone en el estudio de Noriega y Pourrahimian (2022) [14], donde se exploran mejoras en la asignación de recursos y equipos, lo que permite un seguimiento más eficiente de los activos operacionales. Estos avances incluyen la optimización del consumo energético y la gestión eficiente de los equipos y materiales, lo que resulta fundamental para maximizar el rendimiento de las operaciones.

En cuanto a la estabilidad de taludes en minas a cielo abierto, las técnicas de Machine Learning (ML) y Deep Learning (DL) han demostrado ser altamente efectivas. Wang et al. (2023) [15] aplicaron algoritmos de aprendizaje profundo para predecir el coeficiente de estabilidad en taludes, obteniendo una mayor precisión en comparación con los enfoques tradicionales. Estos avances son esenciales para la gestión proactiva de los riesgos geotécnicos y la seguridad de las operaciones mineras.

Por último, Choi et al. (2022) [16] demostraron mediante el uso de modelos híbridos, que la combinación de Harris Hawks Optimization (HHO) con Support Vector Machines (SVM) podría mejorar la eficiencia operativa y aumentar la producción en las minas a cielo abierto, optimizando el desempeño del equipo de transporte. Si bien HHO no está incluido en el estudio actual, podría evaluarse su uso en la optimización de hiperparámetros de modelos predictivos de consumo de potencia o en la optimización de variables operativas en tiempo real para mejorar la eficiencia energética.

Procesamiento de minerales

Los modelos desarrollados en el área del procesamiento de minerales se suelen centrar en predicción de tonelaje en molinos y chancadores, consumo de potencia en molinos, y control de variables en celdas de flotación.

El estudio realizado por Avalos, Kracht y Ortiz (2020) [1] examina el uso intensivo de energía de los molinos SAG en minería, destacando la imprecisión de métodos actuales en la predicción de consumo energético en tiempo real. Propone alternativas utilizando variables operacionales y técnicas de aprendizaje automático y profundo, como regresión polinómica y redes neuronales,

para predecir con mayor eficiencia el consumo energético. El estudio resalta la relevancia de cambios en la matriz energética y cómo la minería se adapta a estos, además de abordar las limitaciones de técnicas actuales y la creciente atención hacia métodos de regresión basados en datos reales para mejorar la eficiencia y sostenibilidad en la minería.

Otro enfoque para integrar la molienda en el procesamiento de minerales ha sido el uso de algoritmos no supervisados. Li et al. (2020) [17] propone una estrategia de aprendizaje automático para optimizar las rutas de procesamiento y reducir la variabilidad en las capacidades de tratamiento a lo largo de la vida útil de la mina. Este modelamiento se realiza mediante la agrupación de bloques de material en función de características como el grado de mineral y su composición. Esta agrupación permite optimizar las rutas de procesamiento, ajustando las capacidades de las plantas de tratamiento según las características del material en cada etapa. Al reducir la variabilidad entre los bloques de material procesado, se garantiza una operación continua y eficiente, minimizando el tiempo de inactividad y maximizando la productividad a lo largo de la vida útil de la mina.

En los estudios más recientes, Saldaña et al. (2023) [18] propusieron la aplicación de ML para mejorar los procesos de molienda en la industria del cobre en Chile, con el objetivo de reducir el consumo energético en molinos SAG y aumentar la recuperación de minerales procesados. Utilizaron una combinación de redes neuronales y árboles de decisión para predecir el consumo energético y anticipar el comportamiento del proceso bajo diferentes condiciones operativas, considerando variables como el tonelaje procesado, la granulometría y la velocidad del molino. Estos modelos permitieron identificar y corregir cuellos de botella operativos al detectar ineficiencias en el proceso, a través del análisis de discrepancias entre las predicciones y los datos reales.

Estay et al. (2023) [20] analizaron los desafíos clave en la aplicación de ML en el procesamiento de minerales, destacando que uno de los principales problemas es la adquisición de datos y la integración de modelos fenomenológicos. Estas limitaciones, combinadas con la complejidad de los procesos de molienda y flotación, sugieren que la adopción de ML debe ir acompañada de un enfoque que asegure la calidad y cantidad de datos disponibles. Este trabajo complementa la discusión sobre la integración de ML en la molienda al ofrecer soluciones para mejorar la precisión de los modelos predictivos.

El uso de visión por computadora -tecnología que permite a las máquinas interpretar y analizar imágenes y videos- y algoritmos de detección de fallas -herramientas que identifican patrones anormales en datos operativos- también ha crecido en el procesamiento de minerales. McCoy y Auret (2019) [19] destacaron cómo estas técnicas mejoran la detección temprana de problemas y facilitan el mantenimiento predictivo, reduciendo los fallos imprevistos y aumentando la disponibilidad de los equipos.

En general, los modelos presentados en los estudios revisados no se han implementado completamente con resultados operativos comprobados, y existe una falta de información respecto a su validación a gran escala. Sin embargo, se reconoce que estos enfoques tienen el potencial de optimizar procesos, reducir el consumo energético y mejorar la recuperación de minerales, además de permitir explorar la implementación de estrategias de mantenimiento predictivo.

3. METODOLOGÍA

3.1 USO DE PYTHON Y LIBRERÍAS

En las áreas de ciencia de datos y aprendizaje de máquina, se utilizan diversos programas para el análisis y modelamiento de la información. Python es uno de los más utilizados debido a su facilidad de uso, sintaxis simple y gran cantidad de librerías, lo que facilita la colaboración entre profesionales de la ciencia de datos.

Antes de iniciar el trabajo sobre la base de datos es necesario familiarizarse con el uso de Python y sus librerías, junto con las configuraciones necesarias para utilizar las librerías de modelamiento para aprendizaje profundo, las cuales incluyen el uso de la unidad de procesamiento gráfico o GPU.

Los códigos son escritos y ejecutados a través de los softwares *Spyder* y *Wingware*.

Dentro de las librerías principales que fueron utilizadas para el análisis, visualización y modelamiento se encuentran:

Numpy. Pandas. Seaborn. Matplotlib. CSV. Time. Skylearn. Tensorflow.

Dado que *tensorflow* es una librería utilizada para modelar gráficamente simulando una red de nodos o capas ocultas, necesita capacidad de procesamiento gráfico multidimensional proporcionada por la GPU, y una forma eficiente de acceder a esto es mediante una plataforma llamada CUDA, la cual es utilizada en los códigos de programación entregados y se convierte en un requisito para utilizar los modelos de aprendizaje profundo.

Los códigos utilizados en la presente memoria serán guardados en un repositorio online de acuerdo con lo indicado en anexos (ANEXO 8.1)

3.2 ENTENDIMIENTO DE LOS MODELOS A UTILIZAR

Si bien no se profundiza en el código de programación y funciones matemáticas de cada uno de los modelos, si se debe tener una noción básica de la forma de funcionamiento de cada uno, por lo que es importante interiorizarse en cada uno de los modelos utilizados.

- Modelo de regresión polinomial
- Modelo vecindad de los k-vecinos más cercanos
- Máquina de vectores de soporte
- Perceptrón multicapa
- Memoria de corto-largo plazo

3.3 GENERACIÓN DE BASE DE DATOS

De forma paralela al aprendizaje e instalación de las herramientas a utilizar se debe generar la base de datos, definiendo horizonte temporal, intervalos de tiempo y variables a utilizar de forma preliminar.

El paso a paso de esta etapa se puede visualizar en Figura 5, sección “Preprocesamiento de base” encerrada en un recuadro azul.

Una vez obtenida la base de datos se debe trabajar sobre la limpieza de esta, identificando celdas vacías, con error de sensores y valores fuera de rango, y definir el tratamiento que se les entrega a estos según su tipo y cantidad. Además, se debe realizar reemplazo de las variables cualitativas por valores numéricos, facilitando el análisis computacional.

Sobre la base de datos limpia se realiza un análisis estadístico de la información, estudiando el comportamiento de las variables operacionales en el tiempo, definiendo dominios de estimación mediante variables cualitativas para determinar estados estacionarios y correlaciones lineales entre variables cuantitativas.

Junto con esto se realiza un breve análisis del comportamiento de las variables manipulables y de control de molienda SAG, para definir su variabilidad y rangos de operación.

Para analizar la influencia del estado de revestimiento en el desempeño del modelamiento se debe identificar fechas de cambio de este, agregando un contador de tratamiento de tonelaje durante su vida útil, volviendo a cero cada vez que estos sean cambiados.

A partir de los mantenimientos con cambio de revestimiento mayores definidos se debe separar la base de datos en diversos subconjuntos para realizar las predicciones, generando una subdivisión entre bases de datos para entrenamiento y validación. Para esto, se identifican las fechas donde se realiza el cambio de revestimiento total del molino SAG, situación que altera el peso y volumen interno del molino, junto con modificar la altura del levante de la carga al interior de este.

Para incluir la influencia del estado de revestimiento en los modelos, se agrega un contador de tratamiento como nueva variable, el cual es reiniciado cada vez que se da inicio a una nueva campaña.

A partir de esta separación, se define también una base de entrenamiento general que incluye todos los subconjuntos de bases de entrenamiento y de igual forma con la base de modelamiento.

Según lo indicado en los alcances, los modelos son realizadas utilizando solamente base de entrenamiento y validación, quedando fuera de los alcances el uso de base de prueba.

La cantidad de base de datos a utilizar corresponde a la cantidad de separaciones que se realice a la base completa de acuerdo a los mantenimientos realizados en el período, más las bases compilatorias de entrenamiento y validación.

3.4 MODELAMIENTO

El paso a paso de esta etapa se puede visualizar en Figura 5, sección “Modelamiento” encerrada en un recuadro verde.

Para la realización de los modelos las bases deben ser adaptadas para ser utilizadas con el código entregado. De acuerdo a metodología propuesta en artículo Ávalos et al., (2020) [1] se deben generar nuevas variables para los valores de salida o *target* de acuerdo a los soportes temporales a utilizar, para posteriormente realizar la normalización de base de datos.

La normalización se realiza de forma independiente para cada base de datos.

Luego, y de acuerdo a la metodología a implementar, se agregan variables de variación de velocidad de giro y tonelaje alimentado con el fin de capturar tendencias locales, las cuales deben

ser generadas posterior a la normalización de los datos. Se eligen las variables velocidad de giro y tonelaje de alimentación ya que son variables de control por parte del operador, permitiendo capturar tendencias que podrían impactar el consumo energético.

Una vez que se realiza la adaptación de la base de datos, se debe realizar modificaciones a los códigos de acuerdo a lo siguiente:

- Se debe actualizar el promedio y desviación estándar de la base de datos de entrenamiento a utilizar en el modelamiento. Esto debe variar cada vez que se utiliza una base de entrenamiento distinta.
- Se debe indicar cuál o cuáles serán las variables de entrenamiento a considerar como predictores y el soporte temporal de consumo energético.
- Se debe indicar para cada modelo, de forma independiente, los rangos de variación de sus parámetros de acuerdo a lo siguiente:
 - Regresión polinomial: Variación grados $n = [1, 2, 3 \dots 6]$
 - Vecindad k-vecinos más cercanos: Variación de vecinos $k = [1, 2, 3 \dots 50]$
 - Máquina de vectores de soporte: Variación constante $C = [10^{-3}, 10^{-2} \dots 10^3]$
 - Perceptrón multicapa: Número de capas ocultas $nH = [4, 8, 12 \dots 200]$
 - Memoria de corto-largo plazo: Número de capas ocultas $nH = [4, 8, 12 \dots 200]$

No se realiza modificación de hiperparámetros en modelos de percepción multicapa y memoria de corto-largo plazo, dado que se asume que el ajuste realizado originalmente aplica al caso desarrollado en la presente memoria. Esto ya que ambas bases de datos describen el mismo fenómeno en contextos similares, y serán utilizadas para predecir la misma variable (consumo energético).

Para cada modelo de entrenamiento, se realizan múltiples predicciones separadas en 3 casos, los cuales son aplicados para cada uno de los modelos a utilizar.

Caso 1:

Se utilizan bases de datos separadas por set de mantenimiento / campaña de revestimientos. Anteriormente se definió cuáles serían consideradas como bases de entrenamiento y cuáles como bases de validación, por lo que se deben analizar todas las combinaciones posibles de estas bases.

De esta forma se busca identificar si las campañas se comportan de forma similar al ser utilizadas para entrenar los modelos de forma independiente. También se busca identificar si existe alguna influencia que afecte a los modelos cada vez que ocurre un cambio de revestimiento, comparando el desempeño de los modelos al ser entrenados de forma independiente con cada Set y al ser entrenados de forma global (Caso 2). También se busca saber si el desempeño de la predicción se comporta de forma similar al inicio y final de cada campaña de revestimientos.

Considerando que se tienen 'x' Set i de entrenamiento e 'y' Set i de validación, entonces la cantidad de predicciones a realizar para cada modelo de predicción será 'x*y'.

Caso 2:

Se utiliza el entrenamiento considerando la base de datos completa, es decir, una predicción con base de entrenamiento vs base de validación global. Para esta modelamiento, se analiza el desempeño del entrenamiento considerando nuevas variables de entrada. Esto quiere decir que se realiza el entrenamiento utilizando las variables de la metodología original, y posteriormente agregando nuevas variables que podrían cambiar el desempeño del entrenamiento, como por ejemplo, la influencia de estados de revestimiento o granulometría de alimentación.

También se analiza el desempeño de los modelos al utilizar como base de entrenamiento la utilizada en Avalos et al., (2020) [1] con las variables originales del estudio.

Caso 3:

Finalmente, se generan predicciones variando los soportes temporales utilizados. Es decir, se revisa el desempeño de los modelos al aumentar la ventana de tiempo de la predicción. Para esto, y dado los tiempos de ejecución de la programación, se seleccionan los sets de mantenimiento de menor tamaño para ser utilizados en la predicción de consumo de potencia en todos los soportes temporales considerados.

El objetivo de este caso es analizar el cambio en los desempeños del modelamiento al aumentar la ventana de predicción de los modelos, utilizando target de predicción a 5 minutos, 10 minutos, 20 minutos, 30 minutos y 1 hora.

3.5 EVALUACIÓN DE MODELOS Y COMPARACIÓN DE RESULTADOS

Para cada uno de los casos se procede a realizar la evaluación de desempeño del modelo haciendo uso de la base de validación correspondiente. Para esto, dentro del mismo código de programación, se utiliza el modelo entrenado para predecir los valores de energía consumida de la base de validación, y se compara con el valor real para cada tiempo t_i , entregándose como resultado los valores de error relativo medio $rRMSE$ y coeficiente de correlación lineal CC . Las fórmulas utilizadas para estos parámetros corresponden a:

$$rRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\bar{y}} * 100$$

$$CC = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}}$$

Donde y_i es el valor real, \hat{y}_i es el valor predicho, \bar{y} es la media de los valores reales, $\bar{\hat{y}}$ es la media de los valores predichos y n es el número de datos.

Una vez obtenido la totalidad de resultados, se procede a realizar el análisis de estos. Esto corresponde a los dos últimos recuadros de diagrama de flujo ilustrado en Figura 5.

3.6 DIAGRAMA DE FLUJO

La metodología empleada en este estudio se resume en el diagrama de flujo presentado en la Figura 5:

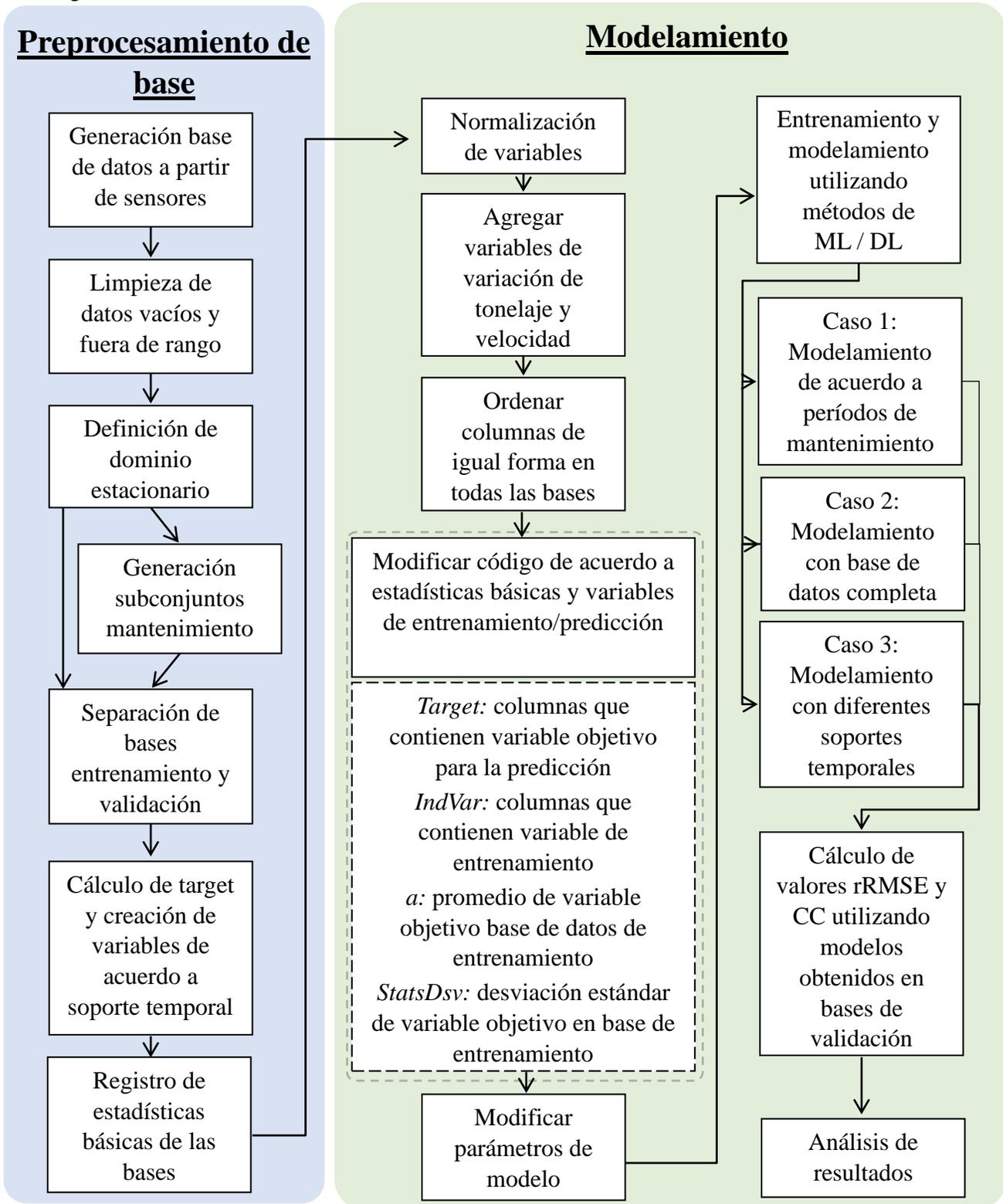


FIGURA 5. DIAGRAMA DE FLUJO METODOLOGÍA

4. APLICACIÓN DE MODELOS

4.1 CASO DE ESTUDIO

Características del circuito

La base de datos a trabajar durante el estudio consiste en datos de operación de un molino SAG en circuito SABC-A, es decir, un molino SAG conectado en circuito indirecto con, en este caso, dos molinos de bolas, y un chancador de pebbles. La disposición de los equipos en un circuito de molienda SAG SABC-A se ilustra en la Figura 6.

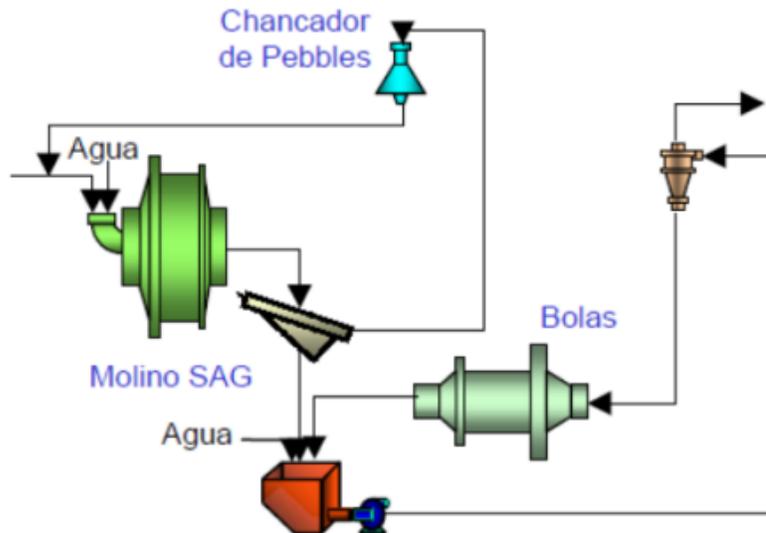


FIGURA 6. DIAGRAMA CIRCUITO MOLIENDA SAG SABC-A

TABLA 1. CARACTERÍSTICAS DE EQUIPOS BASE DE DATOS UTILIZADA

Equipo	Molino SAG	Molino de bolas (2)
Diámetro	40'	26'
Largo	22'	38'
Potencia nominal motor	23.000 [kW]	17.200 [kW]
Velocidad de giro	9-9,8 [rpm]	12 [rpm]
Velocidad crítica	74-80%	68%
Tamaño de bolas	139 [mm]	63-75 [mm]
Rendimiento	4.000 [tph]	1.400 [tph]

Características del control automático en faena

La operación desde la sala de control para el control de la molienda SAG se basa en mantener la utilización de potencia al máximo posible, evitando sobrecarga del molino. Para esto se utilizan diversos sistemas de control que permiten mantener las variables operacionales dentro de rangos definidos y de esta forma alcanzar la potencia requerida.

Para el sistema de control utilizado en molino SAG de datos entregados, se realiza la diferenciación de variables controladas, manipuladas y perturbaciones del sistema.

- Variables controladas: Peso [ton] / Presión de descansos [psi], Potencia [kWh] y generación de Pebble [ton]
- Variables manipuladas: Tratamiento [tph], Velocidad de giro [rpm] y Porcentaje de sólidos [%]
- Perturbación medida: Granulometría fina menor a 2”[%]

El funcionamiento de este sistema consiste en la generación de límites altos y bajos de operación para las variables controladas, que de acuerdo a pruebas de funcionalidad y modelamiento metalúrgico permiten que el consumo de potencia se encuentre en su máximo posible sin riesgo de sobrecarga del molino. Para mantener las variables controladas dentro de los rangos definidos, el sistema de control realiza movimientos en las variables manipuladas.

Estos rangos sólo son modificados de acuerdo al desgaste de revestimiento (tiempo) y la afectación de perturbaciones medidas, lo que queda sujeto a la expertiz de la persona encargada de la sala de control.

4.2 GENERACIÓN DE BASE DE DATOS

La información operacional obtenida que constituye la base de datos a utilizar comprende datos de molino SAG en circuito abierto con datos puntuales cada 5 minutos desde el 01-01-2016 al 10-09-2020, incluyéndose variables operacionales e indicadores de funcionamiento de equipos.

Se obtienen 493,620 datos temporales, donde cada intervalo de tiempo t tiene asociado las variables tonelaje alimentado (FT) [tph], set point de tonelaje alimentado (SpFT) [tph], porcentaje de granulometría menor a 2” alimentado a SAG (PS) [%], velocidad de giro (SSp) [rpm], set point de velocidad de giro (SpSSp) [rpm], potencia consumida (EC) [kWh], peso molino (We) [ton], presión de descansos (BPr) [psi], porcentaje de sólidos calculado (SPe) [%], agua alimentada (Wtr) [m³/h], estado de funcionamiento de molino SAG (RunSAG), estado de funcionamiento de molinos de bolas (RunBM1-2) y estado de funcionamiento de bombas de alimentación a baterías de hidrociclones (RunP1-2-3-4).

En datos originales se obtiene también los valores de consumo específico de energía [kWh/ton], pero se decide eliminar esta variable ya que consiste en un cálculo y no una variable operacional propia. Además, esta variable posee períodos de tiempo mayores sin datos registrados, donde casi un cuarto de los datos obtenidos corresponde a valores *missing* o *-99*.

4.2.1 PREPROCESAMIENTO DE BASE DE DATOS

El procesamiento de los datos es una etapa fundamental para trabajar en procesos de aprendizaje de máquina, ya que permite el aumento de la calidad de la base sobre la que se trabajará, además de entregar robustez al proceso de entrenamiento, evitando sobreajustar los modelos (*overfitting*).

Dado que los datos obtenidos provienen de sensores de terreno en planta operativa, se tiene gran cantidad de desviaciones en la base, incluyendo fallas de comunicación, configuración, hardware, cálculo, entre otros, provocando que valores medidos no sean guardados en el sistema y por ende registrados como *missing* en la base de datos. Dentro de los errores señalados se encuentran:

No such host is known. Invalid Float. No Data. Comm Fail. Scan Timeout. Pt Created. Calc Failed. Not Connect. Configure. Bad. Scan Off. I/O Timeout. Equip Fail. Unit Down. Error.

El principal problema de estas ausencias es que impiden que modelos de aprendizaje de máquina sean entrenados ya que el código desarrollado para esta aplicación no acepta valores no-numéricos. Para estos casos se decide reemplazar valor no numérico por -99.

Además de valores missing, considerando que la totalidad de valores de la base deben ser números para ser utilizados en el aprendizaje, se realiza reemplazo de los valores de estado FUNCIONANDO, DETENIDO por 1 y 0 respectivamente.

Las estadísticas básicas de la base de datos se encuentran en la Tabla 1, donde se observa que existen valores vacíos -99 para la totalidad de las variables, y donde los valores máximos se escapan de los valores promedios y cuartil 75%.

TABLA 2A. ESTADÍSTICAS BÁSICAS DE BASE DE DATOS PRE-LIMPIEZA

	FT (tph)	SPFT (tph)	PS (%)	SSp (rpm)	SpSSp (rpm)	EC (kWh)	We (ton)	BPr (psi)	SPe (%)	Wtr (m3/h)
count	493,620	493,620	493,620	493,620	493,620	493,620	493,620	493,620	493,620	493,620
mean	4,012	4,186	53.58	7.69	8.01	19,059	875	1,306	68.81	1,243
std	1,441	1,127	19.14	4.18	2.89	127,715	75	330	27.40	460
min	-99	-99	-99	-99	-99	-795	-99	-99	-99	-99
25%	3,511	3,510	49.65	7.75	7.75	16,284	852	1,357	74.12	1,044
50%	4,558	4,590	56.42	8.52	8.52	18,846	888	1,387	76.58	1,362
75%	4,910	4,906	63.03	9.03	9.03	21,012	918	1,417	78.03	1,538
max	7,500	7,500	99.84	9.81	9.80	9,999,000	1,302	1,887	999	2,000
count -99	176	230	143	502	245	92,568	231	489	6,461	284

TABLA 2B. ESTADÍSTICAS BÁSICAS DE BASE DE DATOS PRE-LIMPIEZA

	RunSAG	RunBM1	RunBM2	RunP1	RunP2	RunP3	RunP4
count	493,620	493,620	493,620	493,620	493,620	493,620	493,620
mean	1	-1	-4	1	1	1	1
std	3	12	21	2	2	2	2
min	-99	-99	-99	-99	-99	-99	-99
25%	1	1	1	0	0	0	0
50%	1	1	1	1	1	1	1
75%	1	1	1	1	1	1	1
max	1	1	1	1	1	1	1
count -99	549	7,725	23,810	153	153	165	165

Para identificar valores fuera de rango, *outliers*, se utilizan gráficos de tiempo e histogramas. Estos valores deben ser identificados y tratados ya que distorsionan la distribución de datos interfiriendo en el proceso de aprendizaje (Figura 7 y Figura 8).

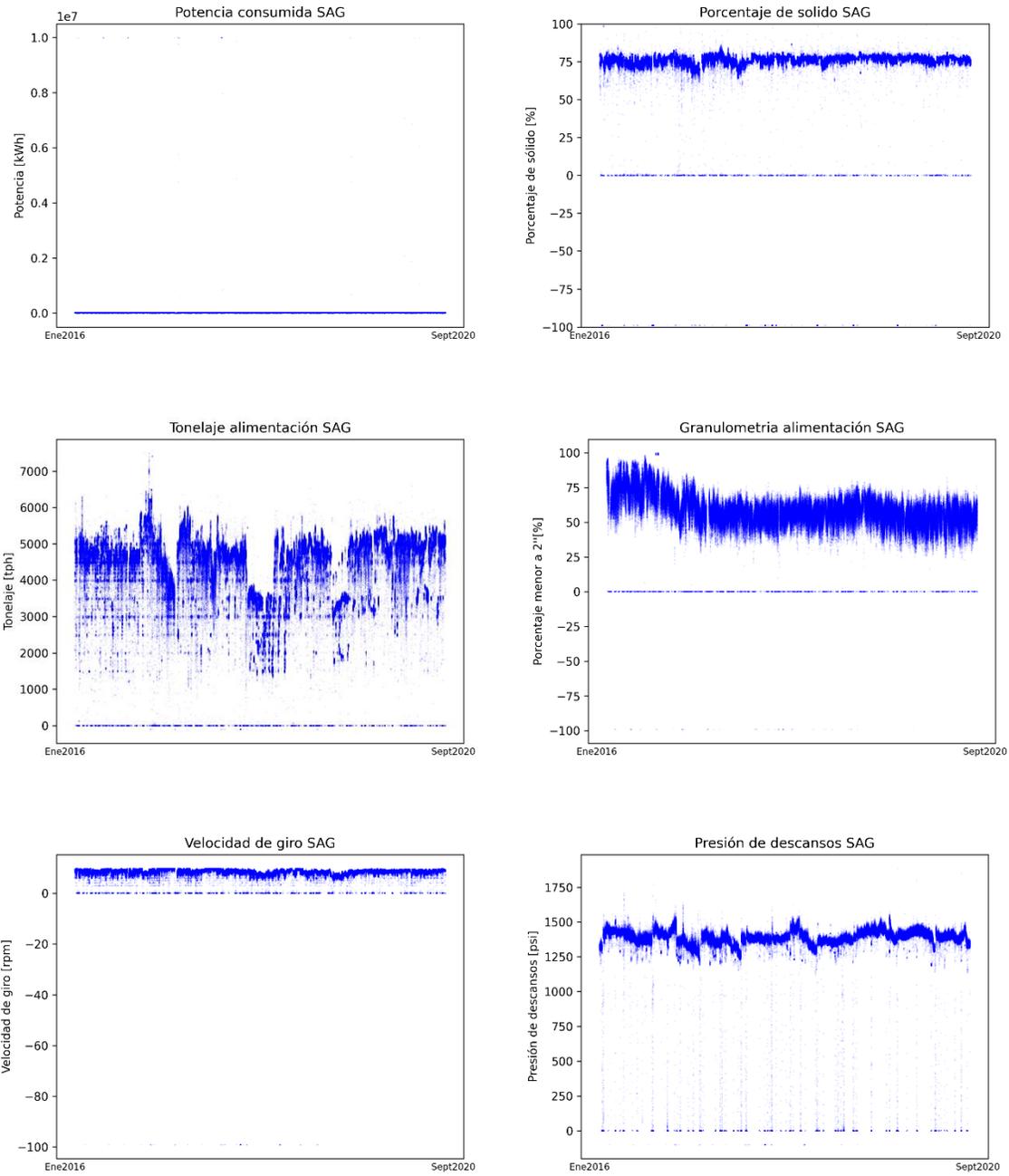


FIGURA 7. VARIABLES PRINCIPALES EN EL TIEMPO PARA BASE DE DATOS PRE-LIMPIEZA PARTE 1

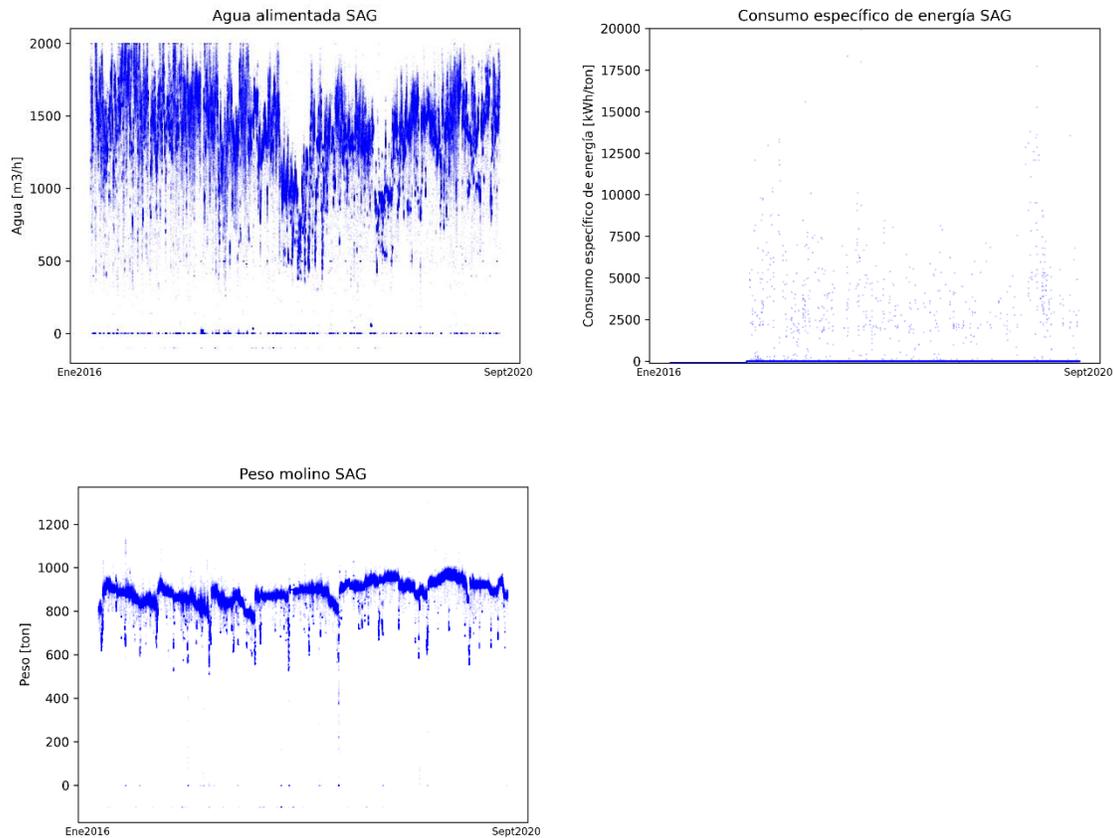


FIGURA 8. VARIABLES PRINCIPALES EN EL TIEMPO PARA BASE DE DATOS PRE-LIMPIEZA PARTE 2

Dada la revisión de estadísticas básicas y gráficos de comportamiento de las variables, se decide dejar fuera de utilización la variable de consumo específico de energía por su gran presencia de valores *-99* y *outliers*, y de esta forma no reducir significativamente la base de datos durante la limpieza. La información contenida en esta variable se respalda al utilizar la potencia consumida y tonelaje alimentado por separado.

Se identifican los siguientes rangos de trabajo para la eliminación de *outliers*.

- Potencia SAG < 25.000 kWh
- Peso de molino > 500 toneladas
- Sólidos 0-100%
- Granulometría 0-100%

Finalmente, para los datos que no pertenecen a los rangos indicados se decide eliminar el registro completo asociado a la medición temporal. Los pasos de limpieza de datos siguen el orden de la siguiente tabla (Tabla 2), generándose una eliminación por etapas donde se muestra la cantidad de filas afectadas y su representatividad con respecto a la base de datos original. Se debe considerar que en la eliminación no se consideran columnas con estados de funcionamiento.

TABLA 3. CONDICIONES LIMPIEZA DE DATOS

Criterio	Datos afectados	
Eliminación valores <i>missing</i> -99	6,956	1,41%
Potencia SAG menor a 25.000 kWh	54	0,01%
Peso mayor a 500 toneladas	143	0,03%
Granulometría 0-100%	0	0,00%
Porcentaje de sólido 100%	30	0,01%
Valores restantes	486,437	98,5%

Para el trabajo con estados de funcionamiento de equipos, se crean dos nuevas variables para agrupar estos valores: cantidad de molinos de bolas funcionando (BMn) y cantidad de bombas alimentación a baterías de hidrociclón funcionando (Pn). Para estos casos, los valores indicados como -99, se considerarán como equipo detenido.

En Tabla 4 y Tabla 5 se registran los valores estadísticos básicos para la base de datos limpia. No se observa presencia de valores *outliers*, mínimos negativos o valores *missing*.

TABLA 4. ESTADÍSTICAS BÁSICAS DE BASE DE DATOS POST-LIMPIEZA PARTE 1

	FT (tph)	SPFT (tph)	PS (%)	SSp (rpm)	SpSSp (rpm)	EC (kWh)	We (ton)	BPr (psi)	SPe (%)	Wtr (m3/h)
count	486,437	486,437	486,437	486,437	486,437	486,437	486,437	486,437	486,437	486,437
mean	4,067	4,225	54.35	7.9	8.13	17,604	878	1,323	71.05	1,260
std	1,372	1,072	18.05	2.26	1.54	5,405	63	296	19.46	440
min	0	0	0	0	0	0	511	0	0	0
25%	3,535	3,596	50	7.79	7.8	16,405	854	1,358	74.24	1,068
50%	4,571	4,593	56.58	8.53	8.53	18,912	889	1,388	76.64	1,366
75%	4,916	4,910	63.13	9.04	9.05	21,040	918	1,418	78.04	1,541
max	7,500	7,500	99.84	9.81	9.8	24,428	1,135	1,887	100	2,000

TABLA 5. ESTADÍSTICAS BÁSICAS DE BASE DE DATOS POST-LIMPIEZA PARTE 2

	RunSAG	BMn	Pn
count	486,437	486,437	486,437
mean	1	2	3
std	0	1	1
min	0	0	0
25%	1	2	2
50%	1	2	3
75%	1	2	3
max	1	2	4

Para analizar el comportamiento de los valores se realiza una revisión a su correlación lineal (Figura 9).

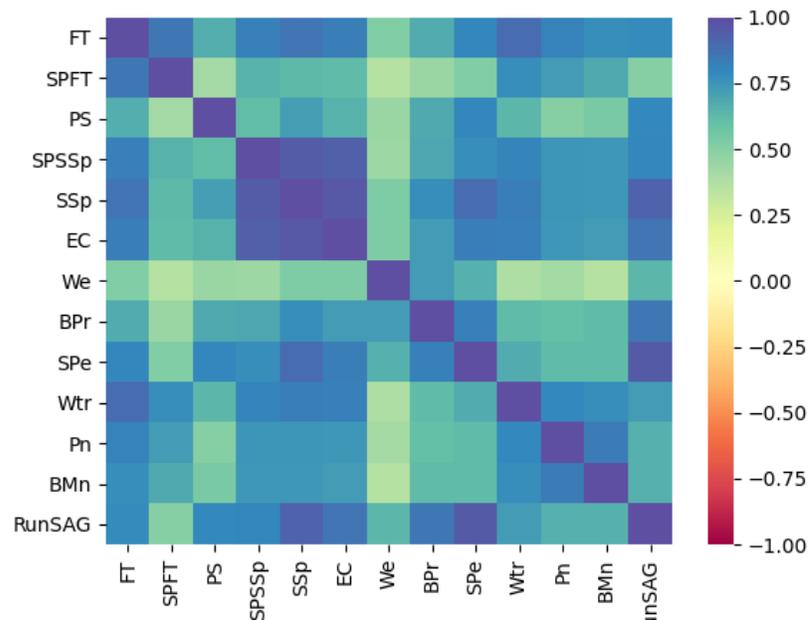


FIGURA 9. MATRIZ DE CORRELACIÓN PARA BASE DE DATOS POST-LIMPIEZA

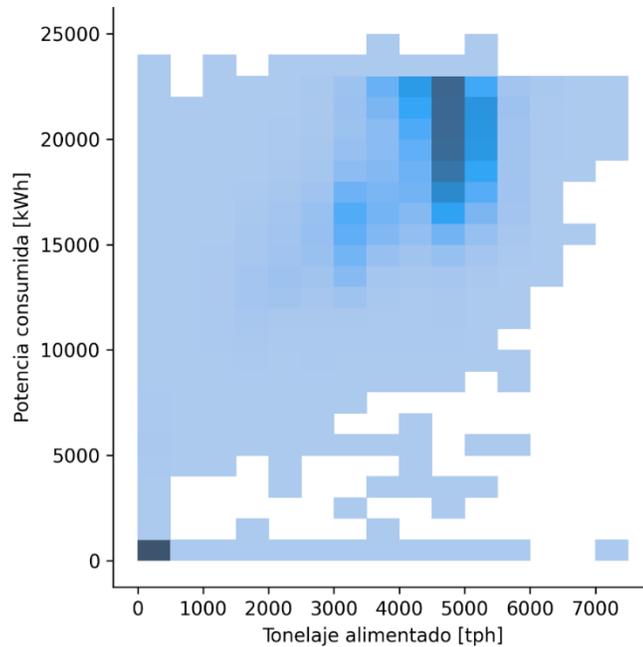
Se observa una alta correlación lineal del consumo energético del molino SAG [EC] con la velocidad de giro [SSp] ($r=0.97$), y medianamente alta con tonelaje alimentado [FT] ($r=0.83$), porcentaje de sólido [SPe] ($r=0.83$) y consumo de agua [Wtr] ($r=0.83$).

Al igual que en estudios previos, no se observa una alta correlación lineal con los valores asociados a carga del molino [We, BPr], a pesar de ser parte de muchos modelos empíricos y teóricos de predicción de consumo energético.

La correlación con el tamaño de partículas [PS] también es baja, lo que indica que el consumo energético no se relaciona linealmente con esta variable, pero al igual que con el resto de las variables, puede existir otro tipo de correlación que no se refleja linealmente.

Estacionariedad de los datos

El objetivo principal del modelamiento basado en aprendizaje de máquina es la captura y reproducción de patrones en la base de datos de entrenamiento, por lo que es necesario asegurar un comportamiento similar en toda la base.



**FIGURA 10. ESTIMACIÓN DE DENSIDAD:
CONSUMO ENERGÉTICO VS TONELAJE ALIMENTADO SAG**

La base de datos actual indica una distribución bimodal del tonelaje alimentado (FT) versus el consumo energético [EC] (ver Figura 10), lo que implica la existencia de dos comportamientos distintos del molino dependiendo del tonelaje alimentado.

Dado que en base de datos obtenida se adquiere información del funcionamiento de los equipos, es posible definir un estado estacionario del sistema, evitando ruidos operacionales que modifiquen la distribución multivariable, alterando el entrenamiento de los métodos. De esta forma se busca identificar también el origen de la perturbación en la alimentación del molino para la separación de los datos y lograr la distribución unimodal.

Para esto se analiza cómo impacta en la potencia consumida el funcionamiento de equipos aguas abajo. Se presentan gráficos de consumo energético versus cantidad de molinos de bolas funcionando y cantidad de bombas utilizadas. Además, se presenta tabla de estadísticas básicas para cada uno de los casos anteriores.

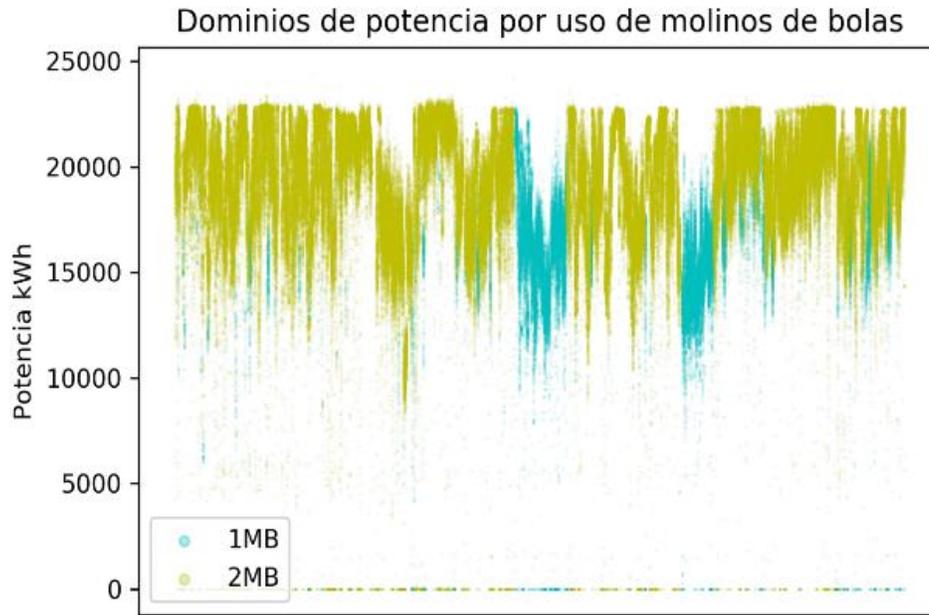


FIGURA 11. DOMINIOS DE POTENCIA SEGÚN NÚMERO MOLINOS DE BOLAS UTILIZADOS

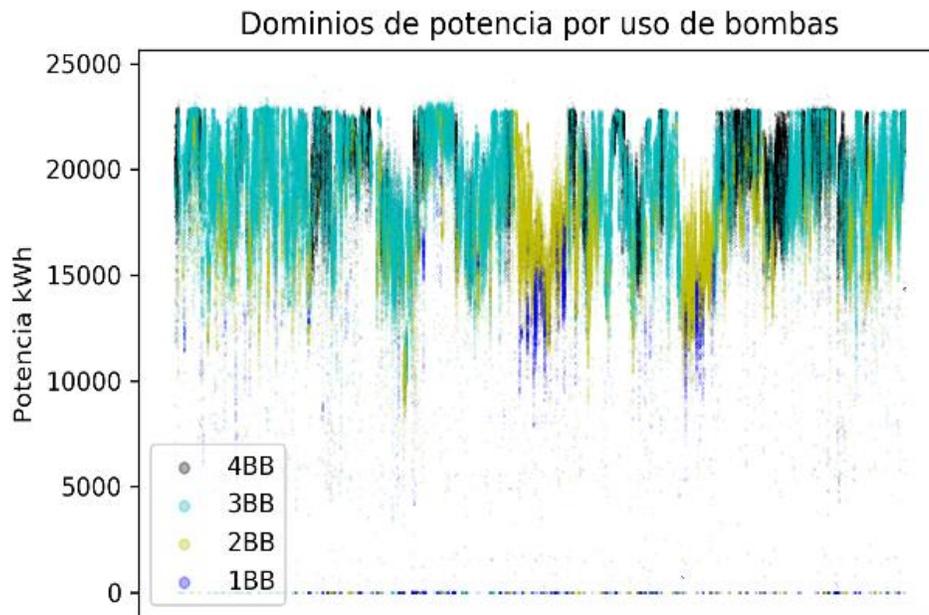


FIGURA 12. DOMINIOS DE POTENCIA SEGÚN NÚMERO DE BOMBAS UTILIZADOS

TABLA 6. ESTADÍSTICAS BÁSICAS DE ACUERDO A DOMINIO

	EC	Uso de bombas				Uso molino de bolas	
	(kWh)	1BB	2BB	3BB	4BB	1MB	2MB
count	486,437	10,918	87,524	252,769	111,640	79,666	383,810
		2.2%	18.0%	52.0%	23.0%	16.4%	78.9%
mean	17,604	9,530	15,507	19,208	20,125	14,894	19,220
std	5,405	6,346	3,997	2,834	2,276	4,447	3,202
min	0	0	0	0	0	0	0
25%	16,405	0	14,474	17,684	18,942	13,969	17,787
50%	18,912	12,418	16,075	19,547	20,551	15,747	19,752
75%	21,040	14,187	17,492	21,306	21,818	17,152	21,445
max	24,428	21,707	23,237	24,147	24,428	23,237	24,428

En la Figura 11 podemos observar que existe una disminución significativa del consumo energético cuando se hace uso de un solo molino de bolas, esto debido a que existe una restricción del tonelaje procesado en el circuito haciendo que molino SAG deba trabajar por debajo de su capacidad instalada. En estos mismos períodos se muestra que se trabaja con 1 o 2 bombas solamente, debido a la misma restricción mencionada anteriormente.

Con respecto al uso de bombas, la mitad del tiempo (52%) se trabajó con 3 bombas, un cuarto del tiempo (23%) con 4 bombas, el resto del tiempo con 2 bombas, salvo un 2% utilizando solo una bomba. También vemos que existe un aumento en el consumo energético a medida que se aumenta el número de bombas de circuito, debido probablemente al aumento de tratamiento del molino. Esto quiere decir que la cantidad de bombas utilizadas tiene una relación directa con el consumo energético en SAG, y por lo tanto con el tonelaje alimentado a este (Figura 10), causando distintos modos de operación.

El 78,9% de los datos pertenece al estado de funcionamiento de 2 molinos de bolas, por lo que corresponde a la operación predominante del sistema. Así mismo, el 75% de los datos corresponde al uso de 3 o 4 bombas de alimentación. (Tabla 4)

Finalmente, para el rango de funcionamiento de 2 molinos de bolas, notamos que la potencia consumida se acerca más a la potencia instalada del molino SAG, por lo que se designará a esto como estado estacionario, dejando fuera de la base de datos la información asociada al uso de un solo molino de bolas.

TABLA 7. EST. BÁSICAS DE ACUERDO AL USO DE BOMBAS PARA EL FUNCIONAMIENTO DE 2 MOLINOS DE BOLAS

	EC	Uso de bombas			
		1BB	2BB	3BB	4BB
count	383,809	498	19,537	252,166	111,444
		0.1%	5.1%	65.7%	29.0%
mean	19,220	5,349	14,560	19,217	20,132
std	3,202	6,478	5,823	2,823	2,259
min	-	-	-	-	-
25%	17,787	-	13,621	17,695	18,948
50%	19,752	0	16,109	19,555	20,554
75%	21,445	12,471	17,986	21,310	21,819
max	24,428	21,170	23,080	24,147	24,428

En la Tabla 5, notamos que de la información restante sólo el 5.2% de los datos corresponde al uso de 1 o 2 bombas, por lo que se decide mantener parte del ruido del circuito y no se realizará limpieza del uso de bombas, en el otro 95% de los datos se mantiene una potencia cercana a la instalada con el uso de 3 a 4 bombas.

Para efectos de la estacionariedad sólo se considerará el funcionamiento de los equipos mencionados, pero existen otras variables que podrían restringir el circuito y que no es posible identificar con los datos actuales (restricciones de agua, restricciones de reactivos, pérdidas de eficiencia por fallas mecánicas, entre otros).

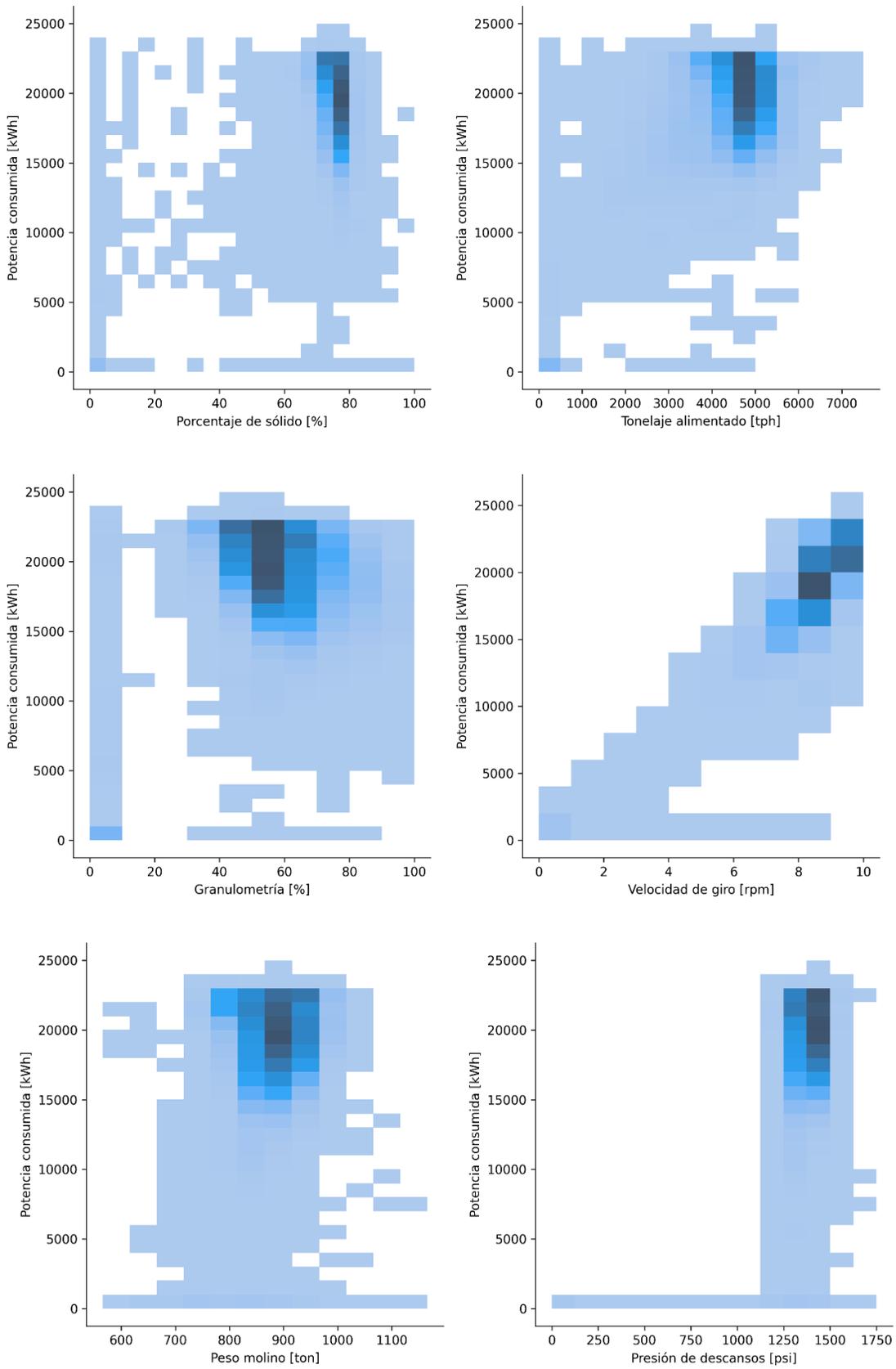


FIGURA 13. ESTIMACIÓN DE DENSIDAD BIVARIABLE PARA BASE DE DATOS DOMINIO 2MB

Finalmente para la base de datos obtenida el comportamiento de las variables principales versus el consumo energético [EC] posee una distribución unimodal en todos sus casos (ver Figura 13), lo que indica que la data se puede normalizar sin perder su representatividad.

4.1.2 SEPARACIÓN POR SETS DE MANTENIMIENTO

Para identificar cómo influye el estado de revestimientos del molino en la predicción del consumo energético se identifican períodos de mantenimientos mayores donde se realiza cambio completo de revestimientos y se separa la base de datos original en 9 sub conjunto de datos (Sets).

TABLA 8. CAMPAÑAS DE REVESTIMIENTO

	Fecha inicio	Fecha fin
Set 1	12-01-2016	30-08-2016
Set 2	30-08-2016	07-04-2017
Set 3	07-04-2017	07-11-2017
Set 4	07-11-2017	07-03-2018
Set 5	07-03-2018	02-10-2018
Set 6	02-10-2018	19-03-2019
Set 7	19-03-2019	08-10-2019
Set 8	08-10-2019	02-04-2020
Set 9	02-04-2020	06-10-2020

Cada uno de estos sets comprende un período cercano a 6 meses (ver Tabla 6), donde los sets iniciales tienen una duración mayor a los sets finales. Esto se puede deber a diferencias en las características del mineral tratado o en incrementos del tonelaje tratado que provoquen una disminución en el intervalo temporal entre cambios de revestimiento.

TABLA 9. ESTADÍSTICAS BÁSICAS POR CAMPAÑA DE REVESTIMIENTOS

Set	count	mean	std	min	25%	50%	75%	max
Set 1	59,240 (15%)	19,457	3,203	-	18,225	20,045	21,580	23,935
Set 2	58,009 (15%)	19,664	3,301	-	18,422	20,544	21,788	24,428
Set 3	53,154 (14%)	18,560	3,998	-	16,558	19,168	21,720	23,503
Set 4	30,838 (8%)	18,588	2,997	-	17,167	18,798	20,462	24,147
Set 5	22,320 (6%)	18,517	3,092	-	17,213	18,932	20,386	23,614
Set 6	44,572 (12%)	18,761	3,501	-	17,137	19,095	21,399	23,446
Set 7	29,538 (8%)	19,919	2,526	-	18,910	20,371	21,537	23,318
Set 8	46,191 (12%)	19,854	2,403	-	18,635	20,113	21,448	23,441
Set 9	37,029 (10%)	19,269	2,288	-	17,880	19,419	20,963	22,972
Data completa	380,891 (100%)	19,217	3,204	-	17,779	19,752	21,447	24,428

La Tabla 7 presenta las estadísticas principales asociadas a cada set, donde se observan comportamientos similares en cuanto al promedio de consumo energético, pero variado en cuanto a la distribución estándar.

A pesar de que los períodos de tiempo son similares, la cantidad de datos para cada conjunto difiere debido a la eliminación de valores faltantes y atípicos en la etapa de preprocesamiento.

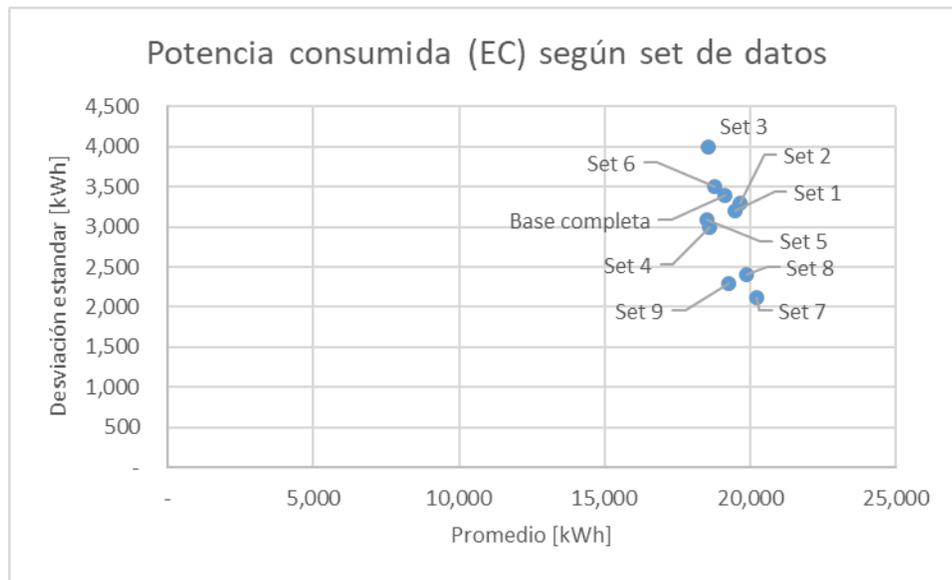


FIGURA 14. PROMEDIO VERSUS DESVIACIÓN ESTÁNDAR SEGÚN SET DE DATOS

La Figura 14 muestra lo indicado anteriormente, donde los valores estadísticos principales de cada conjunto de datos, específicamente la desviación estándar y el promedio de consumo de potencia, se encuentran en la vecindad de los valores característicos de la base de datos completa. Se observa que los sets 6, 7, 8 y 9 tienen una desviación estándar de los valores de potencia mucho menor en comparación con los sets iniciales, lo que indica una menor variabilidad de los datos de fechas más recientes. Esto debe ser tomado en consideración para la separación de bases en el entrenamiento de los modelos.

Dado que las características de cada set no son iguales, se deben registrar los diferentes valores para ser utilizados en la normalización y desnormalización de las variables operacionales.

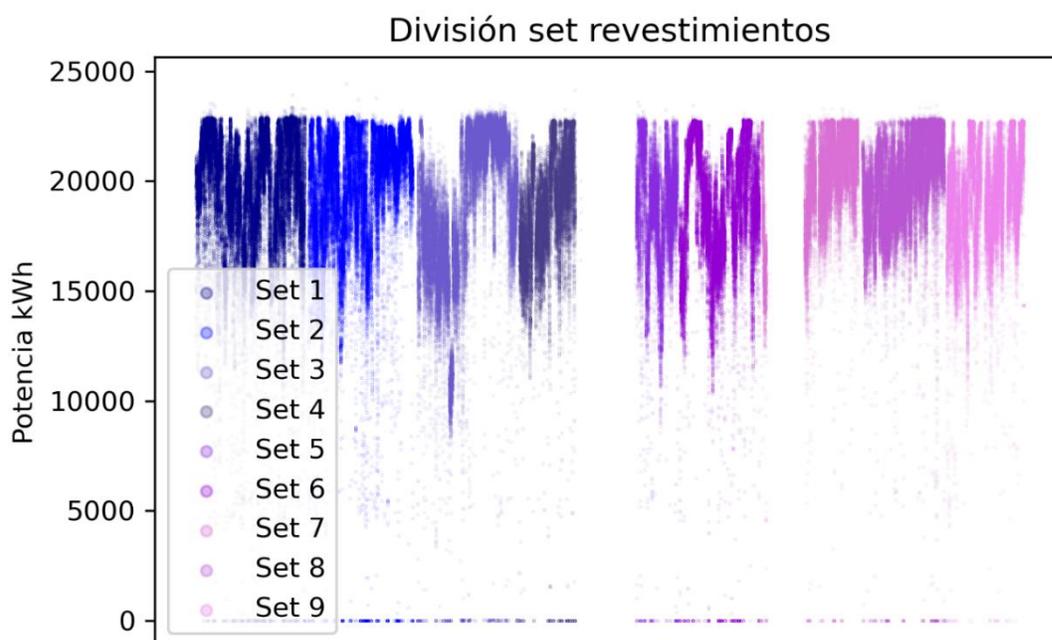


FIGURA 15. CONSUMO DE POTENCIA EN EL TIEMPO SEGÚN SET DE DATOS

Finalmente, en la distribución temporal de los datos se observa que existen 2 períodos con una eliminación masiva de datos. El primer período entre Set 4 y 5, y el segundo en medio del Set 7. En este caso se decide eliminar data inicial del set 7 para asegurar que el comportamiento sea consistente en el tiempo. Se eliminan 2,675 datos del set, quedando con 26,863 (7,1%) datos.

Separación de bases

Con la base de datos revisada, definimos la separación para entrenamiento y validación. Como se menciona en la metodología, no se realizará separación para testeo.

TABLA 10. DEFINICIÓN SET DE ENTRENAMIENTO Y VALIDACIÓN

	Fecha inicio	Fecha fin	Datos	Tipo	Separación base completa
Set 1	12-01-2016	30-08-2016	59,240	Entrenamiento	294,996 (78%)
Set 2	30-08-2016	07-04-2017	58,009		
Set 3	07-04-2017	07-11-2017	53,154		
Set 4	07-11-2017	07-03-2018	30,838		
Set 5	07-03-2018	02-10-2018	22,320		
Set 6	02-10-2018	19-03-2019	44,572		
Set 7	19-03-2019	08-10-2019	26,863		
Set 8	08-10-2019	02-04-2020	46,191	Validación	83,220 (22%)
Set 9	02-04-2020	06-10-2020	37,029		

Tal como se presenta en la Tabla 8, se designan de los sets 1 al 7 para el entrenamiento de los modelos, y sets 8-9 para realizar la validación. Esta separación está hecha para incluir los distintos tipos de variabilidad en los datos, donde del set 6 al 9 existe un comportamiento más estable de la data, decidiéndose dejar set 6 y 7 como parte del entrenamiento.

Esto nos lleva a dejar la separación en 78%/22% entrenamiento/validación.

Para incluir la variable de desgaste de revestimientos se agrega una nueva variable de tonelaje acumulado para cada set (FTR [ton]). Es decir para cada tiempo t_i se agrega el tonelaje procesado desde la fecha de inicio de la campaña correspondiente $t_o^{(Set)}$.

Para finalizar la designación de sets se separan archivos por cada uno enumerándolo del 1 al 7 (Set 1,2...7 respectivamente), y archivos compilatorios de la data completa utilizada para entrenamiento y validación (*Base de entrenamiento* y *Base de validación* respectivamente).

Predicción consumo energético

Para la predicción del consumo energético, se debe contar con una variable de salida o *target* el cual será el valor que debe predecir nuestro modelo. Para este estudio la variable a predecir corresponde al consumo energético [EC].

La base de datos posee valores medidos para cada variable en tiempos t_i espaciados cada 5 minutos, en particular el consumo energético. Como el objetivo es predecir el consumo energético futuro, se deben crear nuevas variables correspondientes a un valor que nos indique el consumo energético EC_{i+n} a predecir utilizando soporte temporal de tiempo n .

Para esto, se definirá el consumo energético en una ventana de tiempo n como el promedio simple de los consumos desde el tiempo t_{i+1} hasta t_{i+n} .

Por ejemplo, para una ventana de tiempo n de 30 minutos en el futuro se define el consumo promedio en los siguientes 30 minutos para un tiempo t_i como:

$$EC_i^{(30min)} = \frac{EC_{i+5min} + EC_{i+10min} + EC_{i+15min} + EC_{i+20min} + EC_{i+25min} + EC_{i+30min}}{6}$$

Este método es utilizado para generar el *target* de predicción de consumo para 5 minutos, 10 minutos, 20 minutos, 30 minutos y 1 hora.

Para cada una de las variables *target* mencionadas, se crea una nueva columna en la base de datos, realizándose un cálculo independiente para cada base de datos (Sets y compilatorias). Así, para cada tiempo t_i se asignan predicciones $EC_i^{(5min)}$, $EC_i^{(10min)}$, $EC_i^{(20min)}$, $EC_i^{(30min)}$ y $EC_i^{(1h)}$.

4.3 MODELAMIENTO

Normalización

Para la realización de los modelos, comenzamos por normalizar la base de datos. Esta normalización utiliza los valores promedio y desviación estándar, lo que la vuelve sensible a valores *outlier* debiendo realizarse luego de limpiar y trabajar la base de datos.

La normalización se realizará de forma independiente en cada archivo de set de datos, además de bases compilatorias.

Para cada variable, y en cada archivo a utilizar, existe una desviación estándar s_{var} y un promedio m_{var} , donde el valor normalizado $x_t^{(norm)}$ para un valor x_t corresponderá a:

$$x_t^{(norm)} = \frac{x_t - m_{var}}{s_{var}}$$

Luego de realizar el entrenamiento de los modelos predictivos, la predicción del consumo energético y_t corresponderá a un número normalizado. Para realizar la transformación inversa se utilizará el promedio y desviación estándar de esta variable de acuerdo con el set de datos de entrenamiento. De esta forma, la transformación inversa quedaría:

$$EC_t^{(pred)} = y_t * s_{EC}^{(train)} + m_{EC}^{(train)}$$

Importante destacar nuevamente que para cada modelamiento realizado la desnormalización de la variable de consumo energético será utilizando los valores característicos de la base de datos de entrenamiento y no de validación. Lo anterior para evitar la distorsión del espacio normalizado, y con ello los patrones capturados por los modelos predictivos entre variables operacionales, al momento de modelar con valores de la base de datos de validación. Los valores a utilizar varían de acuerdo a la base utilizada para entrenar y el soporte temporal de la predicción, variando los *targets* de predicción de 5 minutos hasta 1 hora.

TABLA 11. PROMEDIO CONSUMO ENERGÉTICO SEGÚN BASES DE ENTRENAMIENTO

Promedio	Potencia consumida [kWh]				
	EC_5min	EC_10min	EC_20min	EC_30min	EC_1h
Set 1	19,457	19,457	19,457	19,457	19,456
Set 2	19,664	19,664	19,664	19,664	19,664
Set 3	18,560	18,560	18,560	18,560	18,560
Set 4	18,588	18,587	18,587	18,586	18,585
Set 5	18,517	18,517	18,516	18,515	18,512
Set 6	18,761	18,761	18,761	18,761	18,761
Set 7	20,228	20,228	20,228	20,228	20,228
Entrenamiento	19,114	19,114	19,114	19,114	19,114

TABLA 12. DESVIACIÓN ESTÁNDAR CONSUMO ENERGÉTICO SEGÚN BASES DE ENTRENAMIENTO

Desviación estándar	Potencia consumida [kWh]				
	EC_5min	EC_10min	EC_20min	EC_30min	EC_1h
Set 1	3,203	3,157	3,081	3,019	2,885
Set 2	3,301	3,261	3,198	3,148	3,037
Set 3	3,998	3,967	3,924	3,890	3,817
Set 4	2,997	2,951	2,882	2,828	2,707
Set 5	3,092	3,063	3,017	2,980	2,896
Set 6	3,501	3,476	3,437	3,405	3,326
Set 7	2,126	2,096	2,046	2,003	1,904
Entrenamiento	3,390	3,354	3,298	3,252	3,148

A continuación, en las Tablas 11 y 12, se presenta un resumen de medias y desviaciones estándar en las variables operacionales para cada uno de los Sets de entrenamiento.

TABLA 13. PROMEDIO VARIABLES OPERACIONALES SEGÚN BASES DE ENTRENAMIENTO

	Promedio							
	FT [tph]	PS [%]	SSp [rpm]	EC [kWh]	We [ton]	BPr [psi]	Spe [%]	Wtr [m3/h]
Set 1	4,462	72.88	8.54	19,458	872	1,401	73.56	1,515
Set 2	4,444	60.03	8.66	19,664	856	1,378	73.50	1,478
Set 3	4,485	53.92	8.73	18,560	839	1,352	74.01	1,373
Set 4	4,525	53.06	8.47	18,588	869	1,377	75.36	1,407
Set 5	4,043	53.76	8.51	18,518	861	1,364	75.81	1,221
Set 6	4,486	56.79	8.33	18,761	922	1,359	75.55	1,319
Set 7	4,785	53.59	8.68	20,228	896	1,426	76.80	1,416
Entrenamiento	4,469	59.22	8.56	19,114	872	1,378	74.56	1,409

TABLA 14. DESVIACIÓN ESTÁNDAR VARIABLES OPERACIONALES SEGÚN BASES DE ENTRENAMIENTO

	Desviación estandar							
	FT [tph]	PS [%]	SSp [rpm]	EC [kWh]	We [ton]	BPr [psi]	Spe [%]	Wtr [m3/h]
Set 1	781	13.04	1.17	3,203	33	55	9.46	313
Set 2	999	13.68	1.33	3,301	39	91	10.59	283
Set 3	970	10.40	1.38	3,998	41	107	11.32	321
Set 4	668	8.88	1.06	2,995	14	76	8.89	243
Set 5	1,018	8.99	1.20	3,089	39	83	9.17	310
Set 6	814	10.41	1.27	3,501	20	87	10.36	246
Set 7	543	8.28	0.71	2,126	23	54	5.57	181
Entrenamiento	877	13.44	1.24	3,391	42	88	10.01	295

Predicción

Antes de iniciar el modelamiento, se agregará a cada uno de los archivos generados (*Set 1-9*, *Base de entrenamiento* y *Base de validación*) una columna considerando la variación del tonelaje alimentado y la velocidad de giro entre el tiempo t_{i+1} y t_i , esto dado que al ser variables sujetas a decisión operacional, se plantea encontrar tendencias en su variación de forma local al identificar su variación en el tiempo. El cálculo de estas variables debe ser realizado posterior a la normalización.

De esta forma, las variables base a utilizar corresponderán a tonelaje de alimentación, velocidad de giro, presión de los descansos, variación de tonelaje alimentado y variación de la velocidad de giro.

Las predicciones para cada modelo se dividen en 3 casos:

- **Caso 1** Predicción por set: Se realiza la predicción utilizando cada una de las bases de entrenamiento (*Set 1 a 7*) versus cada una de las bases de modelamiento (*Set 8 y 9*), correspondiendo a un total de 14 modelos por *target* de tiempo. En base a la capacidad computacional, y a los tiempos de entrenamiento de cada método, se decide correr este caso manteniendo *target* y variables operaciones de acuerdo a lo siguiente:
 - *Input*: Variables base
 - *Target*: EC (30 min)

- **Caso 2** Predicción con bases completas: Se realiza la predicción utilizando la totalidad de la *Base de entrenamiento* y *Base de validación*. Se añaden nuevas variables al modelamiento correspondientes a: granulometría menor a 2” en alimentación y tonelaje acumulado por campaña de revestimientos. Las pruebas se realizan entrenando el modelo adicionando cada variable por separado y luego ambas juntas, correspondiendo a un total de 4 modelos de predicción. Además, se decide realizar una prueba utilizando la base de entrenamiento de Ávalos et al., (2020) [1] y la *Base de validación* propia para evaluar los resultados obtenidos. Para esto último se debe considerar adaptar la base de validación propia, realizando un reordenamiento de las columnas input y output y que correspondan los índices con la base del artículo original. Para el modelamiento con esta base de datos distinta, se decide mantener la desviación estándar y promedio de la base original para la normalización inversa. Estas predicciones también se realizarán utilizando como *target* único la predicción del promedio a 30 minutos.
 - *Input*: Variables base + granulometría + campaña de revestimientos (Etto, Etto + FTR, Etto + PS, Etto + FTR + PS)
Base de entrenamiento artículo Ávalos et al., (2020) [1] (EttoReferencia)
 - *Target*: EC (30 min)

- **Caso 3** Predicción de distintos *targets* de tiempo: Se realiza el modelamiento para una misma base de entrenamiento, cambiando únicamente los *targets* de predicción, variándolos entre 5 minutos y 1 hora. En base a la capacidad computacional, y a los tiempos de entrenamiento de cada método, se decide que los modelos se realizarán utilizando únicamente en el Set 7 para el entrenamiento debido a su menor tamaño (26,863 datos). Para la validación se decide utilizar el Set 8.
 - *Input*: Variables base
 - *Target*: EC (5min), EC (10min), EC (20min), EC (30 min), EC (1 hora).

Los modelos obtenidos serán comparadas mediante la obtención de valores de error relativo a la raíz del error cuadrático medio (relative root mean squared error, rRMSE) y coeficiente de correlación (CC) entre los datos predichos mediante el entrenamiento en la base de validación y el valor real de la base.

Resultados

-Regresión polinomial (PR)

Para los modelados de regresión polinomial se decide utilizar variación de grado 1 a 6.

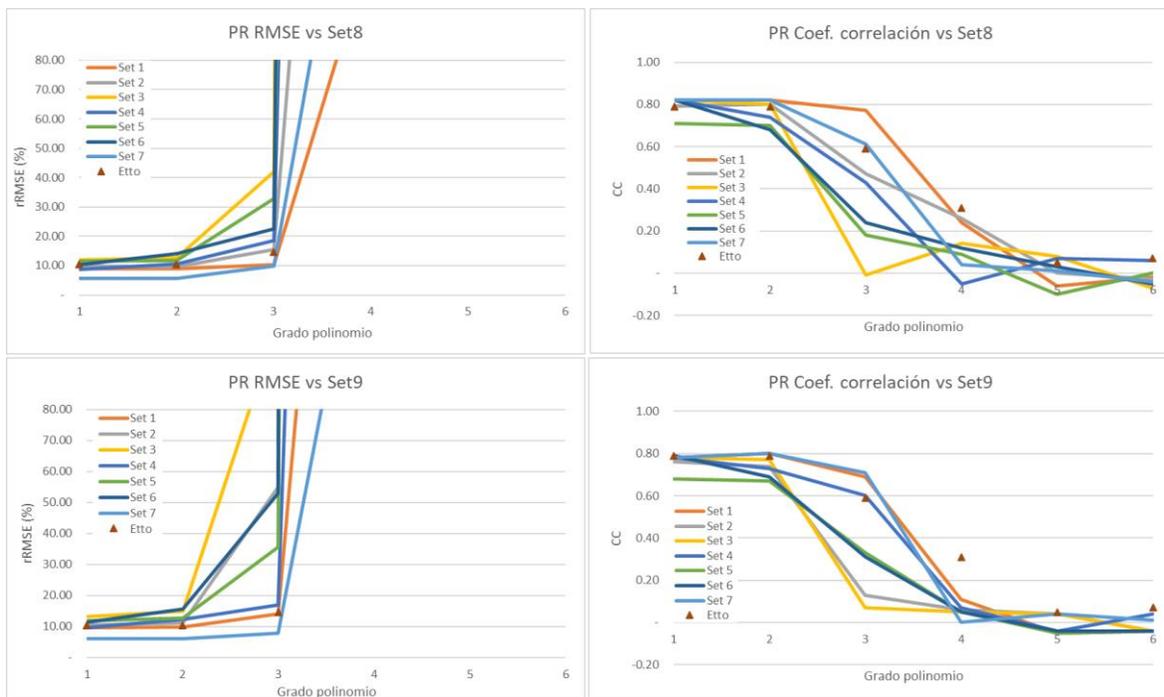


FIGURA 16. MODELAMIENTO CASO 1 REGRESIÓN POLINOMIAL

Para los modelos de caso 1 del método de regresión polinomial se observa que todos los sets tienen comportamientos que siguen la misma tendencia, donde el rRMSE aumenta a medida que se agregan grados al polinomio y el coeficiente de correlación lineal o CC disminuye. El modelamiento por set tiene un mejor desempeño con los grados 1 y 2, y los valores de rRMSE y CC se mantienen en torno a la predicción con bases completas (ver Figura 16).

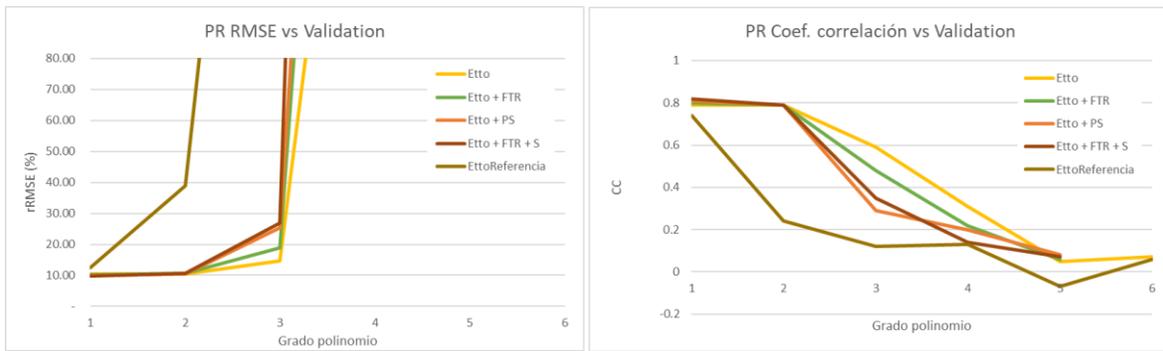


FIGURA 17. MODELAMIENTO CASO 2 REGRESIÓN POLINOMIAL

En los modelados de caso 2 (Figura 17), notamos que también todas las predicciones, excepto con base EttoReferencia tienen un comportamiento similar, donde para grados 1 y 2 los valores de rRMSE y CC son prácticamente idénticos, y a medida que aumenta el grado del polinomio el error también aumenta, siguiendo un comportamiento tipo exponencial.

Al añadir variables como tonelaje por revestimiento [FTR] o tamaño de partícula [PS], el error aumenta y la correlación disminuye desde el grado 3, teniéndose mejores resultados con la base de datos de variables operacionales FT, SSp y BPr.

La forma de las curvas asemeja a las obtenidas en bibliografía[1], sin embargo con la base de datos actual se tiene un error mucho mayor al obtenido en la realización del artículo, donde para grado 4 ya se supera el 80% de error. Esta comparación será discutida en detalle en el siguiente capítulo.

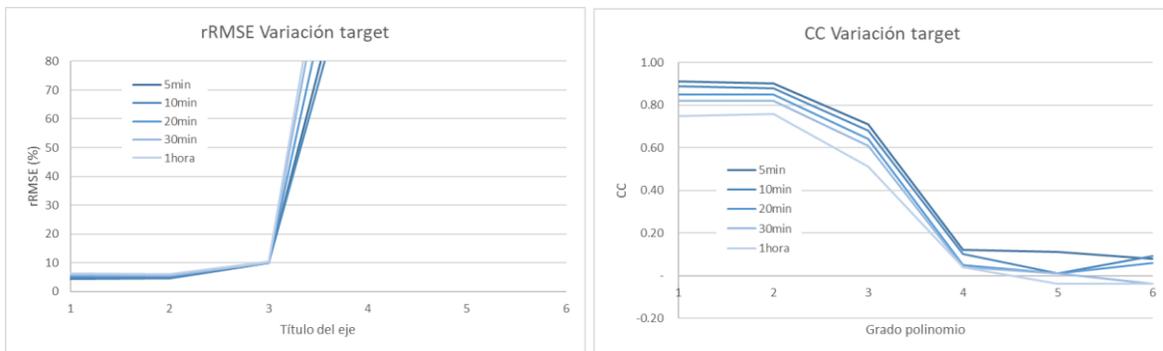


FIGURA 18. MODELAMIENTO CASO 3 REGRESIÓN POLINOMIAL

Como se puede ver en la Figura 18, el error obtenido al modelar con distintos *targets* de tiempo es similar hasta el grado 3 de polinomio, donde para un grado 4 o superior se separan las líneas y aumenta considerablemente el error a medida que la ventana de tiempo es mayor, de forma que el menor error se mantiene a los 5 minutos y el mayor a 1 hora.

En cuanto a la correlación lineal de los resultados, también presenta un resultado esperado, donde los tiempos menores de predicción tienen una mejor correlación, provocando una sobreposición de las líneas hasta llegar a la mínima correlación en el tiempo máximo (1 hora). Para el valor de CC los distintos *targets* siguen el mismo comportamiento a medida que aumenta el grado del polinomio.

-Vecindad de k-vecinos más cercanos (kNNR)

Para los modelados de vecindad de k cercano se decide utilizar variación de vecinos de 1 a 50.

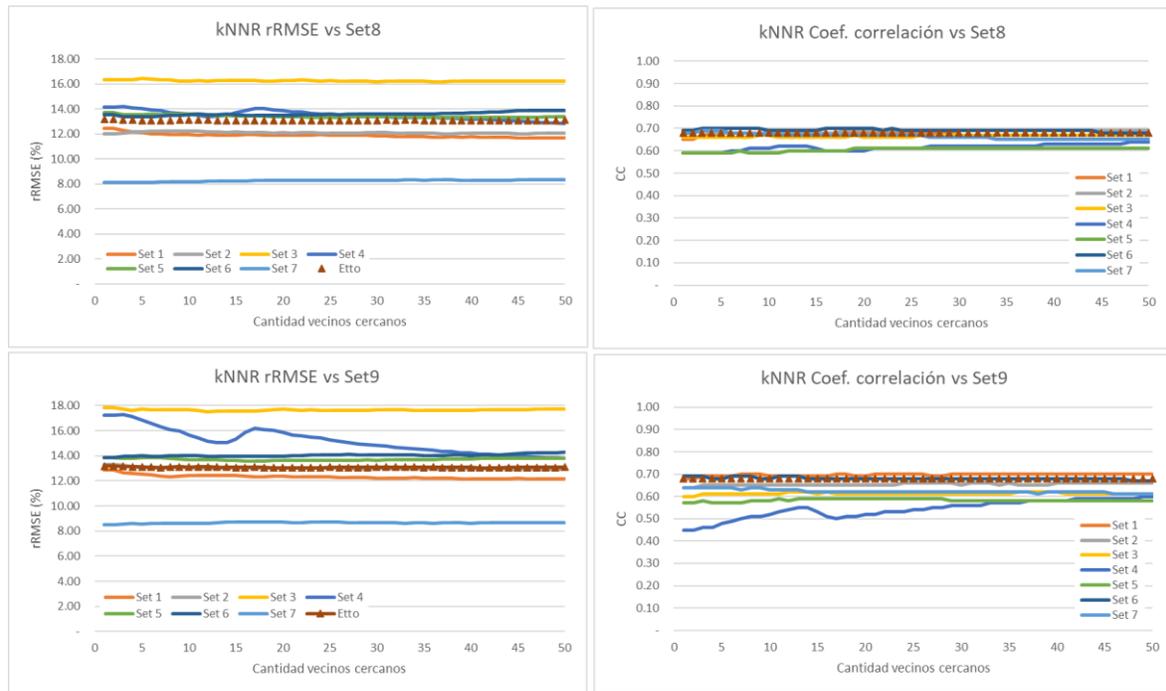


FIGURA 19. MODELAMIENTO CASO 1 VECINDAD K CERCANO

Para los modelados de caso 1 con el método de vecindad de k-vecinos más cercanos (ver Figura 19) notamos que todos los modelos, independiente de la cantidad de vecinos, siguen un patrón horizontal tanto para rRMSE como CC. Esto indica que el modelo no es sensible al número de vecinos que estamos considerando, entregando un resultado similar para 1 y 50 vecinos. A pesar de esto existe un caso aislado de comportamiento de Set7 (entrenamiento) vs Set 9 (validación), donde las curvas de error y correlación presentan un comportamiento errático con respecto al resto de los sets.

Los errores rRMSE generados por este método son muy sensibles a los valores originales de desviación estándar de cada uno de los sets de entrenamiento, donde para ambas bases de validación los errores están ordenados de la mayor desviación estándar a la menor (Set 7 – 3890 kWh y Set 3 – 2003 kWh).

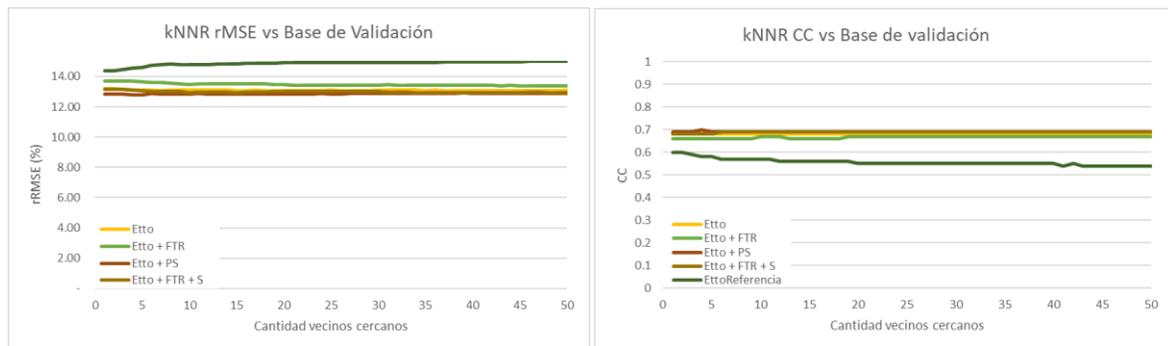


FIGURA 20. MODELAMIENTO CASO 2 VECINDAD K CERCANO

En Figura 20, notamos que en los modelados de caso 2 se observa el mismo comportamiento indicado anteriormente, donde se tiene una curva prácticamente horizontal que no posee sensibilidad en su resultado de acuerdo con el número de vecinos empleado.

Al igual que en predicciones de regresión polinomial, al realizar el modelamiento con base de entrenamiento EttoReferencia [1] se tiene un error ligeramente mayor a los obtenidos con las bases de entrenamiento de este estudio, pero aun así se mantiene en un rango que es incluso mejor que el observado en estudio de la referencia [1].

A pesar de la mínima variación del error al agregar variables a la base original, notamos una leve disminución del error al realizar el modelamiento considerando la granulometría de alimentación, y al contrario un aumento del error al considerar en la base el tonelaje por campaña de revestimiento.

La magnitud de los valores obtenidos tanto para rRMSE como CC asemejan a los obtenidos en el trabajo de Ávalos et al., (2020) [1], alrededor de un 14% de rRMSE (Caso A, Anexo 2) y 0.70 para correlación lineal (Caso A y B, Anexo 2, curva 0.5h)

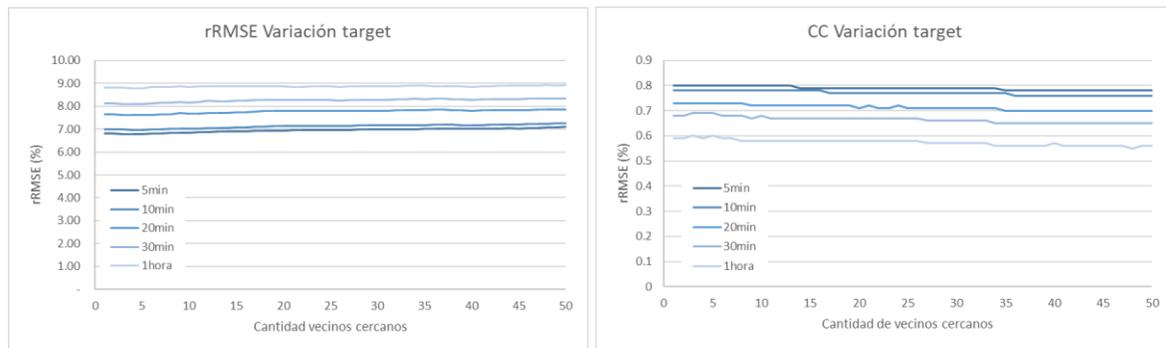


FIGURA 21. MODELAMIENTO CASO 3 VECINDAD K CERCANO

El comportamiento con distintos *targets* de tiempo para la predicción continúa teniendo la misma forma prácticamente horizontal que entregó este modelo para los casos anteriores, manteniéndose una curva de error mínimo con la predicción a muy corto plazo (5 minutos) y aumentando a medida que se agranda la ventana de tiempo (Figura 21).

Al igual que lo observado en caso 2, los valores obtenidos se asemejan en magnitud y forma de curva a los mostrados en Anexo 2.

- Regresión con vectores de soporte (SVR)

Para los modelados usando regresión con vectores de soporte se decide variar el parámetro C de 0.001 a 10,000 con presentación en escala logarítmica.

Por restricciones de capacidad computacional y tiempos de ejecución para cada modelamiento, para caso 2 (modelado con base completa) se disminuye rango de parámetro C, mostrándose de 0.001 a 100.

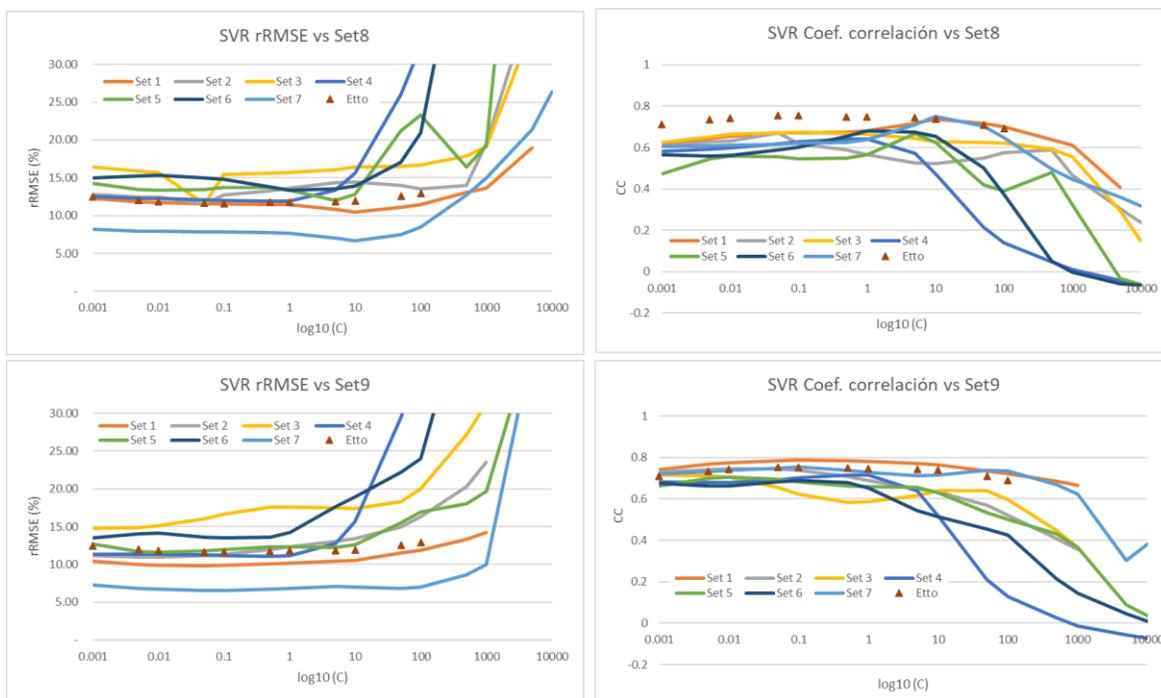


FIGURA 22. MODELAMIENTO CASO 1 REGRESIÓN CON VECTORES DE SOPORTE

Para el modelamiento de caso 1 se mantienen las observaciones anteriores, donde el comportamiento de las curvas obtenidas es similar entre sí, a la vez que mantienen el comportamiento de la curva utilizando Base de Entrenamiento completa (ver Figura 22).

Utilizando ambos sets de validación se mantiene la forma de curva representada en Anexo 3, donde hasta el parámetro $C \sim 100$ se tiene una tendencia sutil convexa en el caso de rRMSE y sutil cóncava en el caso de CC. Para valores de parámetro C mayores, se observa un aumento considerable del error y disminución de la correlación lineal con respecto a lo indicado anteriormente.

Al igual que los casos anteriores, la diferencia de magnitud de los errores es sensible a la desviación estándar indicada para la base de entrenamiento, donde para valores menores de C el menor error se presenta al utilizar el Set7 para entrenamiento (menor desviación estándar), y los errores mayores al utilizar el Set3 para entrenamiento (mayor desviación estándar).

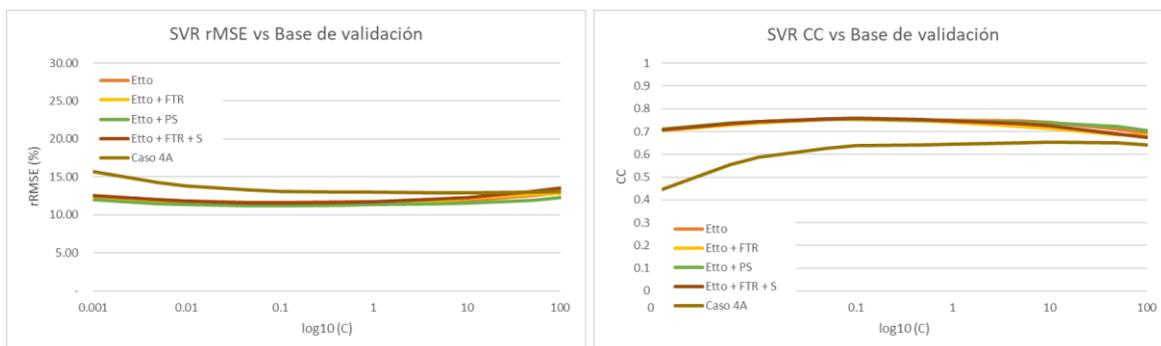


FIGURA 23. MODELAMIENTO CASO 2 REGRESIÓN CON VECTORES DE SOPORTE

Los modelados de caso 2 para el modelamiento por regresión con vectores de soporte, cuyos resultados se indican en Figura 23, muestra una curva que no varía su forma ni su magnitud al agregar las variables de tonelaje por campaña de revestimiento ni tamaño de partícula, manteniendo su error mínimo cercano a 12%, y correlación máxima cercana a 0.75.

Al realizar el entrenamiento con la base EttoReferencia [1] el error es levemente mayor para valores menores de C, acercándose a la curva del estudio actual a medida que el valor de C aumenta.

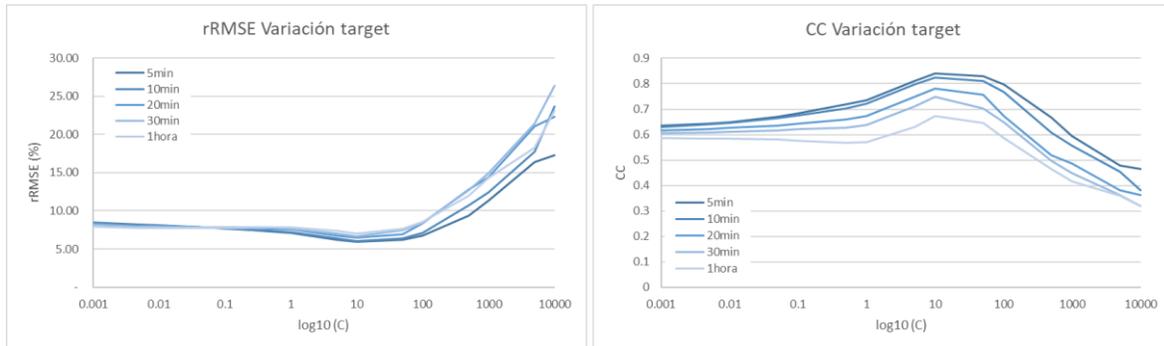


FIGURA 24. MODELAMIENTO CASO 3 REGRESIÓN CON VECTORES DE SOPORTE

Para los modelados con variación en los *targets* de predicción, la Figura 24 muestra cómo se mantiene la forma de curva para todos los horizontes temporales, cambiando la magnitud del error y la correlación lineal de estos, alcanzando un mínimo de error cercano al 7% y máxima correlación de 0.85.

Los errores entregados para todos los tiempos de predicción se mantienen con valores similares hasta $C=0.1$ (rRMSE entre 8% y 9%, CC entre 0.6 y 0.7), punto donde las curvas empiezan a separarse observándose para valores mayores de C un menor error con ventanas de tiempo menores.

Para la correlación lineal se ve que esta aumenta a medida que el valor de C crece, alcanzando su máximo en $C=10$ para todos los horizontes temporales, y posteriormente decreciendo de forma significativa.

-Perceptrón multicapa (MLP)

Para los modelados de perceptrón multicapa se decide variar el número de capas ocultas de 4 a 300.

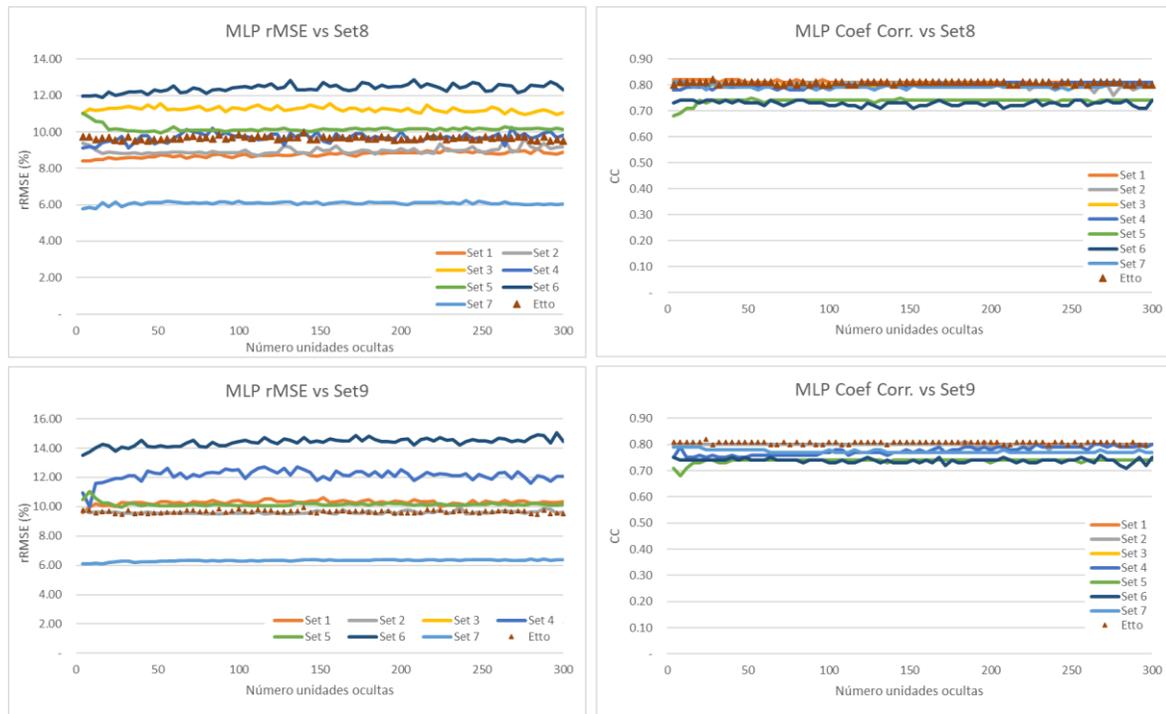


FIGURA 25. MODELAMIENTO CASO 1 PERCEPTRÓN MULTICAPA

Para los modelos por Set de revestimiento (Figura 25) se observa, al igual que con los métodos de Machine learning, que las magnitudes de los errores obtenidos son dependientes de los valores de desviación de la base de entrenamiento, donde para valores mayores de desviación, se obtiene un mayor error. La forma de las curvas se asemeja a lo mostrado en Anexo 4, con curvas prácticamente horizontales para todos los valores de capas ocultas consideradas.

Las magnitudes de los errores y correlación lineal obtenidas con ambas bases de validación son similares entre sí encontrándose errores bajos para set con menor variabilidad (Set 7). Correlación lineal se mantiene entre 0.7 y 0.8 para todos los sets, corriéndose con ambas bases de validación.

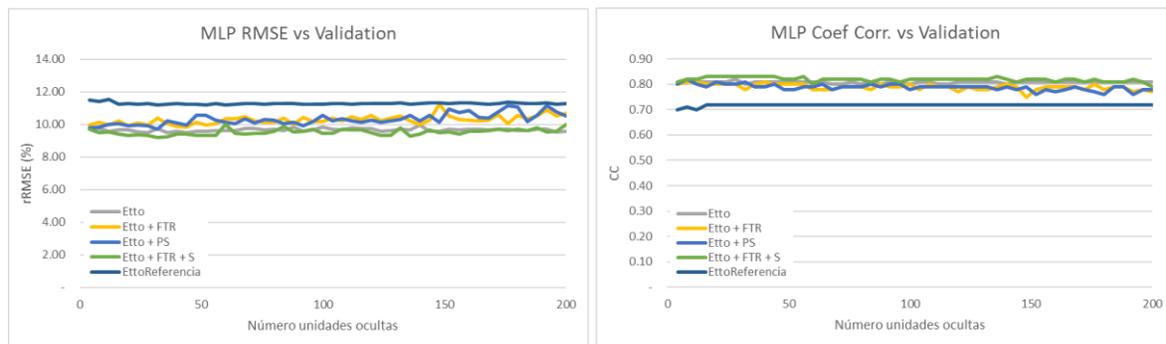


FIGURA 26. MODELAMIENTO CASO 2 PERCEPTRÓN MULTICAPA

A diferencia de los resultados obtenidos con los tres métodos utilizados de Machine learning, para este modelo no existe una mejora de la predicción al agregar las variables operaciones de granulometría y tonelaje por campaña de revestimiento, al contrario el error aumenta y la correlación disminuye levemente para todas las cantidades de capas ocultas consideradas (ver Figura 26).

Al utilizar la base de entrenamiento EttoReferencia se observa un leve aumento del error obtenido subiendo de 10% hasta cerca de 11.5%. Así mismo la variación en la correlación lineal obtenida también es levemente menor, con una disminución de 0.8 a ~0.71.

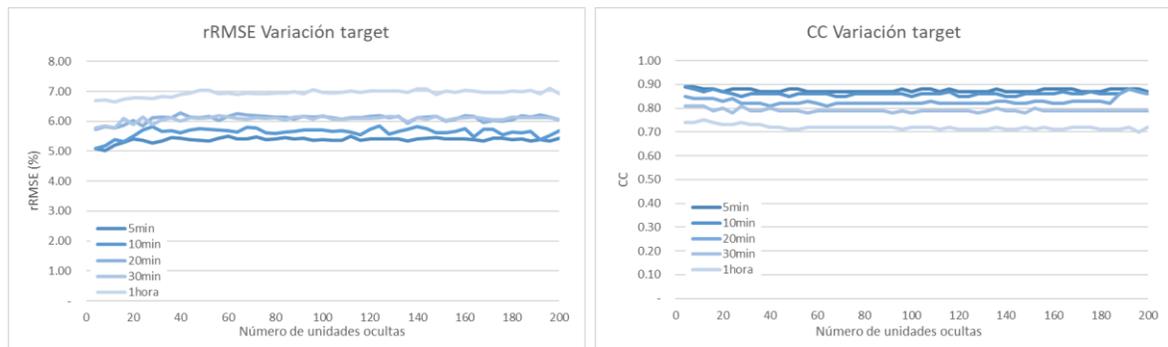


FIGURA 27. MODELAMIENTO CASO 3 PERCEPTRÓN MULTICAPA

En Figura 27, se observa que la variación de *targets* temporales también mantiene su forma horizontal para todos los tiempos considerados, generando un aumento del error y disminución de la correlación a medida que el tiempo de predicción aumenta.

-Memoria corto-largo plazo (LSTM)

Para los modelos de memoria corto-largo plazo se decide variar el número de capas ocultas de 4 a 300.

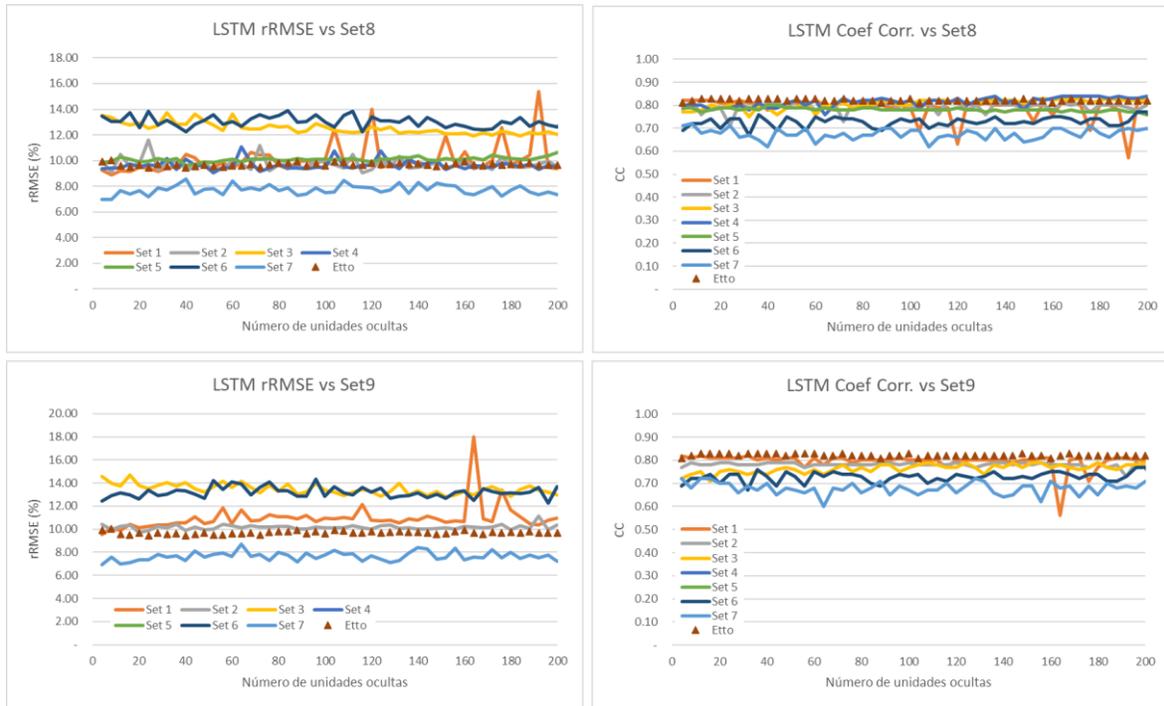


FIGURA 28. MODELAMIENTO CASO 1 MEMORIA CORTO-LARGO PLAZO

Al igual que los resultados ya observados, para el modelamiento utilizando memoria corto-largo plazo los sets presentan comportamientos similares entre sí manteniéndose la forma de la curva en todas los modelos para las diferentes bases de entrenamiento y validación (ver Figura 28).

Sin embargo, a diferencia de los modelos anteriores, el set con menor desviación estándar a pesar de tener el menor error cuadrático también posee la menor correlación lineal (Set 7, Figura 28).

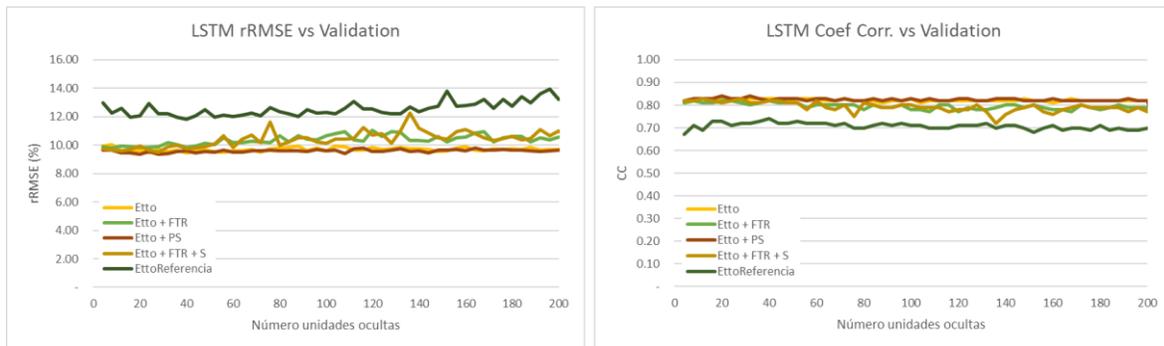


FIGURA 29. MODELAMIENTO CASO 2 MEMORIA CORTO-LARGO PLAZO

Al igual que el método anterior de Deep Learning, perceptrón multicapa, en estos modelados se observa que el desempeño de la predicción disminuye al agregar las nuevas variables, observándose un leve aumento del error cuadrático para todas las cantidades de unidades ocultas, como se muestra en la Figura 29.

A diferencia de los resultados indicados en bibliografía[1] donde el error disminuye al aumentar las capas ocultas de 4 a 300, en Figura 29 esta curva se mantiene horizontal con una leve tendencia al alza al aumentar las unidades ocultas.

En cuanto a la magnitud de los valores, se observa que al utilizar para el entrenamiento la base de EttoReferencia[1] existe un aumento del error del 2% al 4% con respecto al entrenamiento utilizando base.

La correlación lineal se mantiene en los valores indicados en estudio previo[1], siendo cercano a 0.80.

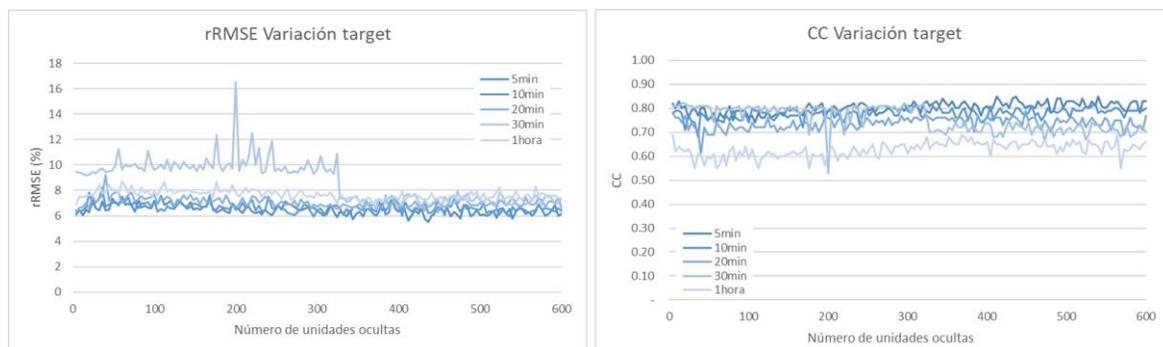


FIGURA 30. MODELAMIENTO CASO 3 MEMORIA CORTO-LARGO PLAZO

Para la variación de los *targets* de predicción de tiempo, lo primero que llama la atención en Figura 30 es la bajada abrupta del error en la predicción promedio de consumo energético en los próximos 30 minutos al llegar a las 328 capas ocultas.

El error cuadrático posterior a las 328 capas ocultas mantiene la tendencia observada con anterioridad, donde la magnitud aumenta levemente al aumentar la ventana temporal de la predicción, y manteniéndose el menor error con la predicción a 5 minutos.

La forma errática de la curva también es coincidente con lo presentado en Anexo 5, así como la magnitud de la correlación lineal a los 30 minutos / 0.5 h.

Comparación de predicciones y valores reales en la validación

Dado que se generaron múltiples gráficas para cada modelo, en este análisis se comparará un nodo representativo de cada modelo utilizando los Sets 1 y 8 para la evaluación de desempeño. A continuación, se presentan las gráficas de dispersión (scatter plot) entre los valores reales y predichos del set de validación para cada modelo (PR, kNNR, SVR, MLP, LSTM).

A continuación, se presentan los valores descriptivos de ambos conjuntos (Set 1 y Set 8), incluyendo las medias y desviaciones estándar de la potencia promedio los próximos 30 minutos, como se observa en la Tabla 15. Las diferencias en las distribuciones de ambos conjuntos afectan los resultados de desnormalización cuando se utilizan las estadísticas de Set 1 para desnormalizar Set 8, tal como es indicado en metodología propuesta basada en bibliografía[1].

TABLA 15. ESTADÍSTICA DESCRIPTIVA PARA NORMALIZACIÓN BASES SET 1 Y SET 8

<i>Target</i> 30 min	Consumo potencia [kWh]	
Método	Promedio	Std Dv
Set 1	19'457	3'019
Set 8	19'853	2'282

Estas diferencias en las medias y desviaciones estándar impactan los valores predichos cuando se utiliza Set 1 para desnormalizar Set 8, lo que ocasiona que algunos valores desnormalizados caigan por debajo de cero, como se muestra en algunos de los gráficos a continuación para diferentes modelos de Machine Learning (ML) y Deep Learning (DL).

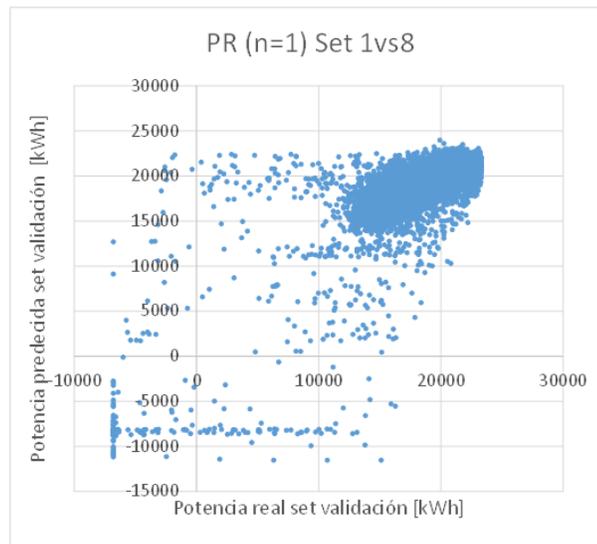


FIGURA 31. GRÁFICO DISPERSIÓN MODELO PR N=1 SET1VS8

El gráfico de regresión polinómica (PR) muestra una dispersión significativa en los valores predichos, con muchos valores negativos y una falta de correlación clara con los valores reales. Esto indica que el modelo presenta dificultades para capturar la relación entre las variables. Los puntos dispersos indican predicciones inconsistentes, lo que se refleja en un CC de 0.79, que indica una correlación moderada.

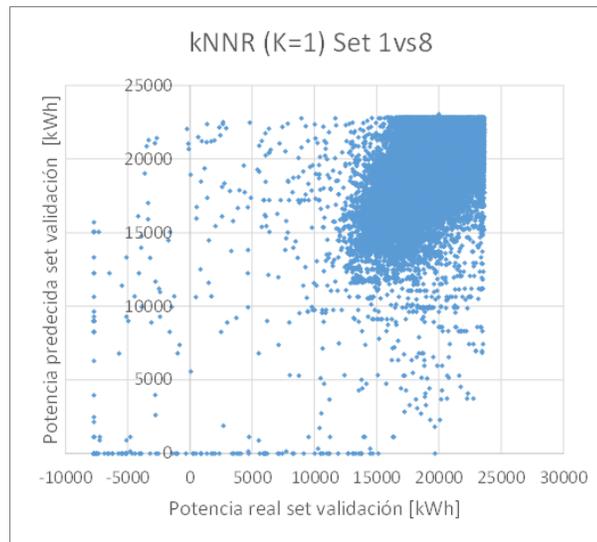


FIGURA 32. GRÁFICO DISPERSIÓN MODELO kNNR K=1 SET1VS8

En el gráfico de kNNR, las predicciones tienden a agruparse en un rango estrecho entre 10,000 y 20,000 kWh, lo que sugiere que el modelo generaliza las predicciones, resultando en menos precisión. El modelo muestra una dispersión significativa, lo que indica una baja correlación. El valor de rRMSE es en general el más alto entre los modelos analizados, lo que confirma que este modelo tiene problemas para capturar correctamente las fluctuaciones en los datos de validación.

Tanto en los gráficos de dispersión de los modelos PR y kNNR, se observan varios valores predichos que se desvían significativamente de la diagonal de predicción ideal (donde los valores predichos y reales deberían coincidir). Estos puntos se identifican como valores anormales o outliers, los cuales generan errores importantes en las predicciones. Esto es probablemente debido a que estos modelos tienen dificultades para capturar patrones complejos en los datos, afectando la precisión de las predicciones

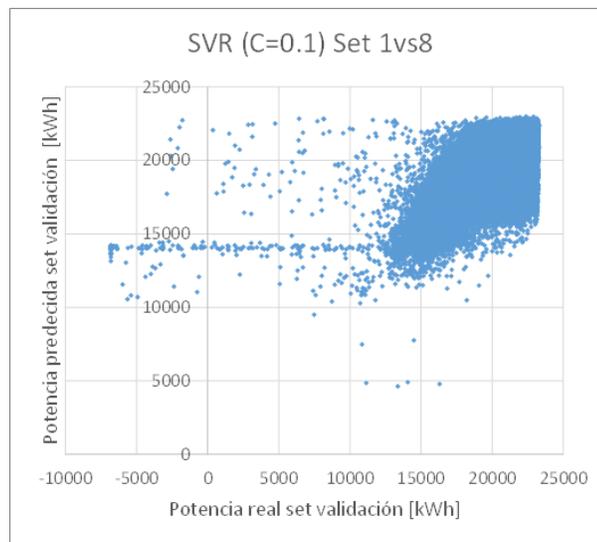


FIGURA 33. GRÁFICO DISPERSIÓN MODELO SVR C=1 SET1VS8

El modelo SVR muestra una correlación moderada, con una alineación más cercana a la diagonal que PR y kNNR. Aunque se observa mayor precisión que PR y kNNR, el modelo aún tiene margen de mejora en comparación con los modelos de Deep Learning.

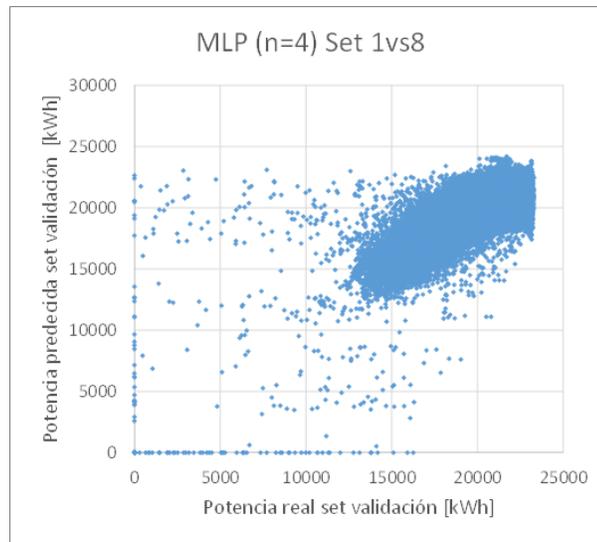


FIGURA 34. GRÁFICO DISPERSIÓN MODELO MLP N=4 SET1VS8

El gráfico de MLP muestra una correlación fuerte, con la mayoría de los puntos alineados cerca de la diagonal, lo que indica un buen ajuste entre las predicciones y los valores reales. El truncamiento identificado en código utilizado para el modelamiento (`np.clip()`) evita los valores negativos, lo que mejora significativamente los resultados de CC y rRMSE calculados en comparación con los modelos de Machine Learning, donde no se observa este truncamiento en los valores.

```
V_r, V_p = int(np.clip(Real * b + a, a_min = 0, a_max = None)), int(np.clip(One_pred * b + a, a_min = 0, a_max = None))
```

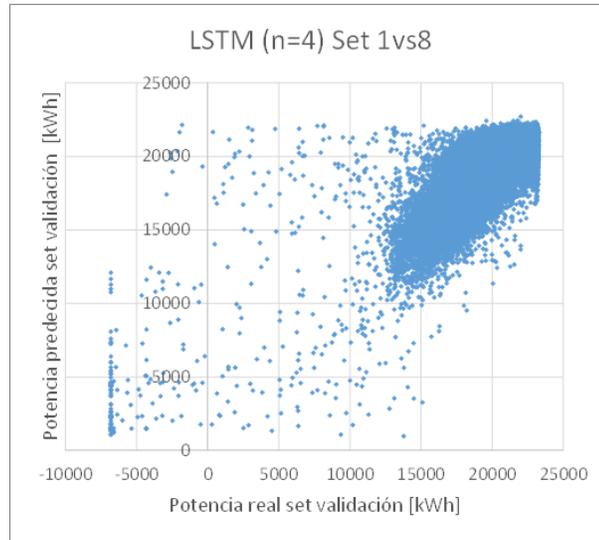


FIGURA 35. GRÁFICO DISPERSIÓN MODELO LSTM N=4 SET1VS8

El gráfico de LSTM muestra la correlación más fuerte entre todos los modelos, con las predicciones alineadas de manera consistente a lo largo de la diagonal y una dispersión mínima. Aunque algunos valores están ligeramente fuera de la diagonal, el modelo presenta un rendimiento sólido. El rRMSE (error relativo cuadrático medio) obtenido para los diferentes nodos es de los más bajos, lo que refuerza la precisión del modelo, similar a lo que ocurre con MLP.

Ambos modelos avanzados, MLP y LSTM, demuestran una capacidad superior para capturar patrones complejos y manejar la variabilidad en los datos operacionales. Esta menor dispersión en los gráficos de predicción confirma que estos modelos son más adecuados para generar predicciones precisas en condiciones variables, destacando su eficacia frente a otros enfoques más simples."

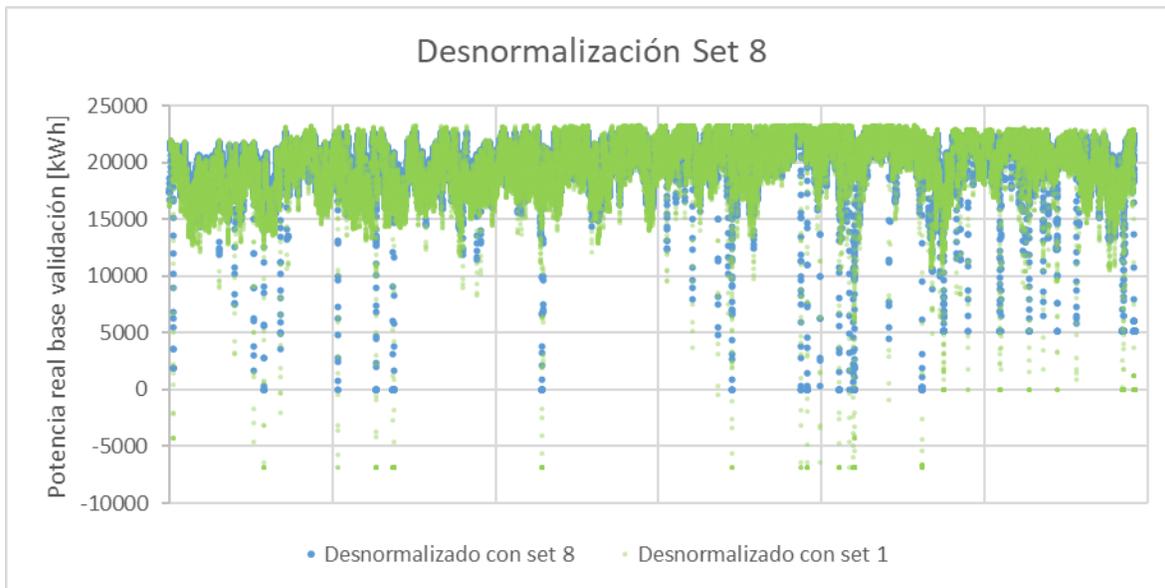


FIGURA 36. DESNORMALIZACIÓN SET 8.

En el gráfico de desnormalización, se observa una comparación directa entre los resultados desnormalizados utilizando las estadísticas del conjunto de entrenamiento (Set 1) y el conjunto de validación (Set 8). La línea verde muestra las predicciones desnormalizadas con Set 1, donde se puede notar que algunas predicciones caen por debajo de cero, lo que genera resultados incoherentes, como valores negativos de potencia. En cambio, la desnormalización utilizando los datos de Set 8 (en azul) muestra resultados más consistentes, sin predicciones negativas. Esto demuestra que las diferencias entre las estadísticas de los conjuntos de entrenamiento y validación afectan significativamente los resultados, lo que debe ser considerado al realizar la desnormalización. En un escenario real, donde solo se cuenta con la base de entrenamiento, esta discrepancia puede generar predicciones fuera de rango, como valores de potencia negativos, que impactan tanto en los valores de CC como en el rRMSE.

5. DISCUSIÓN

Discusión de resultados obtenidos

Dentro de los resultados obtenidos en los modelados por set (caso 1) para los distintos métodos, podemos observar que para todos los modelos la magnitud de las curvas de rRMSE son altamente dependientes de la desviación estándar de la base original, donde las curvas de error para cada set suelen estar ordenadas de acuerdo con su desviación estándar, quedando en caso todos los modelos el Set 7 – 3890 kWh con una magnitud de error que llega a ser prácticamente el doble de la observada en la curva de Set 3 – 2003 kWh.

Esta tendencia se acentúa en los métodos kNNR y MLP, donde al ser curvas horizontales, esta diferencia del doble de la magnitud del error de Set 7 sobre Set 3 se mantiene para todos los parámetros de modelamiento utilizados, a pesar de pertenecer al mismo macro conjunto de datos. Este valor de magnitud llama la atención considerando que la desviación estándar del Set 7 es coincidentemente el doble de la desviación estándar del Set 3 (18% a 9% respectivamente para kNNR, 7,5% a 15% respectivamente para SVR, 6% a 12% respectivamente para MLP y 7% a 14% respectivamente para LSTM).

En base a la observación anterior, se realizan pruebas utilizando la misma base de entrenamiento y modificando la desviación estándar en dos modelamientos distintos. Dada la capacidad computacional se decide correr métodos de machine learning únicamente y utilizar la base de entrenamiento con menor tamaño EttoReferencia [1] y Base de Validación como validación. Se realizan tres modelamientos con target de predicción a 30 minutos, el primero ya ejecutado en caso 2, donde se entrena con EttoReferencia y se considera estadística descriptiva de base propia de este estudio (Tabla 9 y Tabla 10) (Caso 4A), el segundo modelado utilizando estadística descriptiva utilizada en Ávalos et al., (2020) [1] para SAG2 la cual posee una desviación estándar menor ($\bar{x} = 16,710$ y $\sigma = 1,504$) (Caso 4B), y el tercero el cual es realizado sólo para efectos de prueba, se realizará manteniendo la desviación estándar de la bibliografía y el promedio del caso de estudio actual ($\bar{x} = 19,114$ y $\sigma = 1,504$) (Caso 4C).

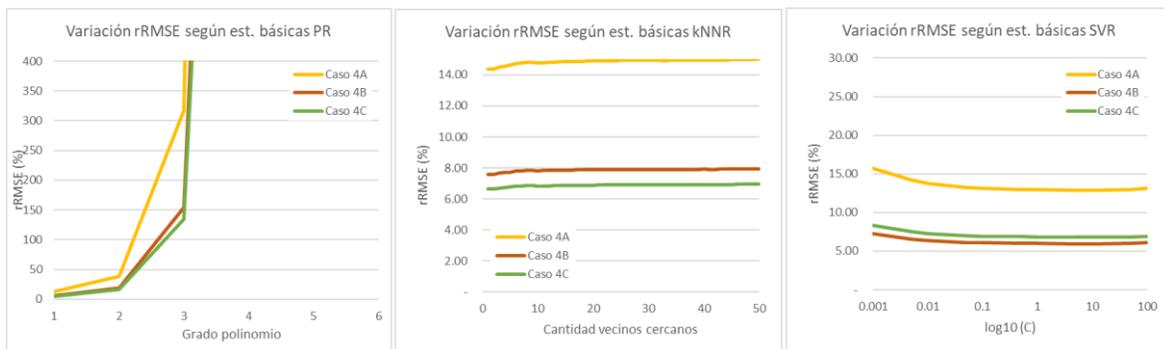


FIGURA 37. VARIACIÓN rRMSE SEGÚN ESTADÍSTICAS BÁSICAS

Lo observado en las gráficas de Figura 31 corrobora lo expuesto, donde para una misma base de entrenamiento y validación, la asignación de una desviación estándar menor disminuye considerablemente el error para todos los parámetros considerados en cada método (Caso 4A vs Caso 4B).

Finalmente esto indica que los errores resultantes de los métodos de modelamiento utilizados son altamente sensibles a las estadísticas descriptivas de la base de entrenamiento, especialmente al valor de la desviación estándar. Esto se debe tener en consideración al comparar resultados cuando se utilizan bases de entrenamiento diferentes.

Para los modelados realizadas con variación del horizonte temporal de la predicción se obtiene que a medida que se aumenta la ventana de predicción (de 5 min hasta 1 hora), existe un aumento del error rRMSE y disminución de la correlación lineal. A pesar de esto, las curvas mantienen su forma entre sí al variar los parámetros para cada modelo.

Las curvas obtenidas para rRMSE y CC para todos los métodos revisados son coincidentes en las formas que se presentaron en bibliografía [1] para sus respectivos métodos, donde las tendencias horizontales, crecientes y convexas se mantienen para el estudio actual. Esto puede deberse ya que ambos modelamientos son realizados para la predicción de consumo energético en molinos SAG, por lo que se puede asumir que los datos persiguen un mismo patrón, teniendo un comportamiento similar a pesar de tratarse de plantas distintas.

Esto también se observa al utilizar la base de entrenamiento EttoReferencia y la Base de Validación propia, las cuales pertenecen a diferentes operaciones sin posibilidad de comprobar si los circuitos o mineralogía son los mismos. En este caso si bien se obtiene un error mayor en los modelos con esta base de entrenamiento, lo cierto es que la curva mantiene la misma forma y la magnitud no supera en más del 30% del error obtenido en el mejor caso para cada método (Ver Figuras 16, 19, 22, 25, 28).

En cuanto a los comportamientos generales observados de las gráficas de variación de rRMSE y CC, para 3 de los 5 modelos revisados estas gráficas permanecen de forma casi horizontal, encontrándose un error y correlación lineal similar para todos los parámetros considerados en los modelos de vecindad k-vecinos más cercanos, perceptrón multicapa y memoria corto largo plazo. Sin embargo, para los modelos de regresión polinomial y regresión con vectores de soporte, existe un crecimiento del error y disminución de la correlación a medida que se aumenta el valor de los parámetros para cada modelo.

En la tabla 15 se presentan para cada método los parámetros con mejor ajuste de acuerdo con lo presentado en las gráficas anteriores de resultados de modelamiento de caso 1, considerando como criterio de selección el error mínimo rRMSE.

TABLA 16. MEJORES RESULTADOS MODELAMIENTO BASE CASO 1

<i>Target</i> 30 min	Machine Learning			Deep Learning	
Método	PR	kNNR	SVR	MLP	LSTM
Modelo	n: 1	K: 19	C: 0.1	n: 28	n: 24
rRMSE	10.37	13.05	11.64	9.50	9.45
CC	0.79	0.68	0.75	0.82	0.83

Primero se realiza una revisión del desempeño de los modelos utilizando las variables operacionales base (FT, SSp y BPr).

En cuanto a los modelos de Machine learning, la mejor predicción corresponde al modelo de regresión polinomial, con un polinomio de grado 1, presentando un error de 10.27% y correlación lineal de 0.79. Así, el desempeño de esta regresión se encuentra por sobre los resultados obtenidos al modelar con métodos de vecindad de k-vecinos más cercanos con un error de 13.05%, el más alto del trío, y regresión con vectores de soporte con un error de 11.64%.

La correlación lineal de la regresión polinomial de grado 1 también es alta, alcanzando un valor de 0.79, por sobre los demás modelos de Machine learning.

A pesar de esto, los mejores resultados se observan con los modelos de Deep learning, donde ambos presentan un error menor y una correlación lineal mayor que los modelos antes revisados.

El menor error, y por ende mejor desempeño, lo tendría el método de memoria corto-largo plazo con 9.45%, y luego con un desempeño similar el método de perceptrón multicapa con un error de 9.50%. Para ambos modelos se tiene una correlación lineal de 0.82 y 0.83 respectivamente, mejor a la observada en métodos de machine learning.

En la tabla 16 se presentan para cada método los parámetros con mejor ajuste considerando el caso 2, donde se añaden variables operacionales. El criterio de selección para el mejor caso corresponde al error mínimo rRMSE.

TABLA 17. MEJORES RESULTADOS MODELAMIENTO CON VARIABLES ADICIONALES CASO 2

<i>Target</i> 30 min	Machine Learning			Deep Learning	
Método	PR	kNNR	SVR	MLP	LSTM
Modelo	n: 1	K: 19	C: 0.1	n: 32	n: 20
Variables adicionales	-FTR -PS	-PS	-FTR	-FTR -PS	-PS
rRMSE	9.71	12.78	11.17	9.22	9.37
CC	0.82	0.70	0.75	0.83	0.84

Con los resultados obtenidos, se observa que todos los modelos con una leve disminución en su error rRMSE, presentándose los siguientes :

- Regresión polinomial: El error disminuyó del 10.27% al 9.71% al agregar las variables de tonelaje por campaña y granulometría de alimentación, manteniéndose como el menor error entre los métodos de Machine Learning.
- k-NN: El error se redujo del 13.05% al 12.78% al considerar la misma cantidad de vecinos pero agregando la variable de granulometría de alimentación.
- SVR: El error disminuyó del 11.64% al 11.17% al incluir la variable de tonelaje por campaña de revestimiento.
- MLP: El error mejoró del 9.50% al 9.22% al incluir las variables de granulometría de alimentación y tonelaje por campaña, modificando el modelo de menor error a MLP.
- LSTM: El error se redujo levemente del 9.45% al 9.37% al añadir la variable de granulometría.

TABLA 18. COMPARACIÓN ERROR MEJOR CASO 1 VS MEJOR CASO 2 POR MODELO

<i>Target</i> 30 min	Machine Learning			Deep Learning	
	Método	PR	kNNR	SVR	MLP
rRMSE Caso 1	10.37	13.05	11.64	9.50	9.45
rRMSE Caso 2	9.71	12.78	11.17	9.22	9.37
Diferencia	0.66	0.27	0.47	0.28	0.08

Aunque todos los métodos empleados muestran una disminución en el error rRMSE (Figura 17), esta reducción es menor al 1% en todos los casos, siendo particularmente menor para los métodos de Deep Learning. No existe una variación significativa en la magnitud de los errores obtenidos para el mejor caso de cada modelo. Esto indica que, aunque los resultados del modelamiento mejoraron, la mejora fue mínima comparada con el aumento del trabajo computacional que implica agregar nuevas variables de entrada.

La falta de una disminución significativa del error rRMSE al agregar variables operacionales sugiere su poca influencia en el entrenamiento de los modelos, especialmente en los de aprendizaje profundo. Esto podría deberse a que estas variables no influyen en el consumo energético o porque ya están implícitamente representadas en las entradas consideradas. Por ejemplo, el tonelaje de campaña de revestimientos, indicativo del desgaste, podría estar contenido en la presión en los descansos (BPr), ya que a medida que aumenta el desgaste de los revestimientos, disminuye el peso del molino. Asimismo, la granulometría puede reflejarse en el tonelaje de alimentación, ya que los cambios en la granulometría afectan los tiempos de residencia del mineral en el molino, lo que a su vez altera el peso del molino. Para controlar estas variaciones, se ajusta el tonelaje de alimentación, aumentándolo cuando los tiempos de residencia en la molienda SAG son menores.

Finalmente, el mejor método de predicción cambió en los nuevos modelamientos al añadir variables operacionales, siendo ahora el modelo de perceptrón multicapa (MLP) con 32 nodos, que incluye las variables base FT, SSp y BPr, y adicionales el tonelaje de revestimiento FTR y granulometría PS. Este modelo obtuvo un error mínimo de 9.22% y una correlación de 0.83

Para realizar una comparación de los mejores modelos indicados en la bibliografía [1] y en el estudio actual con las variables base, se generan las siguientes tablas.

TABLA 19. COMPARACIÓN MODELOS MACHINE LEARNING BIBLIOGRAFÍA CASO 3

	Machine Learning		
	Ávalos et al. (2020) (SAG1)	Ávalos et al. (2020) (SAG2)	Estudio actual
Método	PR	PR	PR
Modelo	n: 2	n: 2	n: 1
rRMSE	8.63	5.87	10.37
CC	0.94	0.87	0.79

TABLA 20. COMPARACIÓN MODELOS DEEP LEARNING BIBLIOGRAFÍA CASO 3

	Deep Learning		
	Ávalos et al., (2020) (SAG1)	Ávalos et al., (2020) (SAG2)	Estudio actual
Método	LSTM	LSTM	LSTM
Modelo	280	596	24
rRMSE	5.61	3.71	9.45
CC	0.94	0.89	0.82

Para ambos estudios encontramos que el mejor método, con el menor rRMSE, es el mismo tanto para Machine learning como Deep learning.

Para el modelamiento mediante aprendizaje de máquina el mejor método de entrenamiento corresponde a la regresión polinomial, donde para el caso de bibliografía el mejor resultado se obtiene con polinomio de grado 2 mientras que para el estudio actual se consigue con un polinomio de grado 1. Como se ha revisado anteriormente la magnitud del error no es una medida exacta de comparación, ya que depende de las estadísticas básicas de la base de entrenamiento, sin embargo podemos ver que es similar entre el caso SAG1 [1] con 8.63% y el estudio actual con 10.37%. La correlación lineal en el estudio actual es menor que el obtenido en la bibliografía, pero aún así alcanza un valor considerable de 0.79.

Para el modelamiento mediante aprendizaje profundo se mantiene como mejor método de las predicciones de memoria corto-largo plazo, alcanzando el error mínimo de todos los métodos realizados tanto en bibliografía [1] como en el estudio actual, con un error de 5.61% y 3.71% en la bibliografía y 9.45% en el estudio actual. Al igual que en modelos de machine learning, la correlación lineal obtenida en el estudio actual es menor a la obtenida en la bibliografía, obteniéndose un valor de 0.82.

A pesar de lo indicado, los valores rRMSE y CC obtenidos son satisfactorios, consiguiéndose en el mejor de los modelos un error de 9.45% y correlación de 0.82 al utilizar las variables base, lo que demuestra que la metodología indicada en la bibliografía es exportable a otros sets de datos para la predicción de consumo de energía en molienda SAG, los cuales pueden tener o no las mismas características mineralógicas, dimensiones y tipo de circuito.

Se debe considerar que los resultados obtenidos son producto de aplicar los códigos de modelamiento diseñados para una planta específica, consiguiéndose un desempeño remarcable al ser utilizados con una base de datos completamente distinta.

Durante el análisis, se identificó que las diferencias en las estadísticas entre el conjunto de entrenamiento (Set 1-7) y el conjunto de validación (Set 8-9) afectan significativamente los resultados de la desnormalización. Por ejemplo, en varios de los modelos al desnormalizar los resultados de Set 8 utilizando las estadísticas de Set 1 se obtuvieron valores incoherentes como potencias negativas, lo que impacta directamente el rendimiento de los modelos en términos de CC y rRMSE. Este fenómeno es importante de considerar, ya que en un escenario real donde solo se cuenta con la base de entrenamiento, esta discrepancia podría generar predicciones que no son

representativas del sistema. Para mitigar este efecto, es esencial tener en cuenta las diferencias en las estadísticas y realizar ajustes adecuados para evitar resultados no realistas.

En el modelo MLP, se implementó un ajuste clave en el código utilizando `np.clip()` para truncar valores negativos en las predicciones y valores reales, evitando así resultados incoherentes. Este ajuste podría contribuir al desempeño del modelo, mostrando valores altos de CC de 0.82 en el mejor de los casos y un rRMSE de 9.50, lo que lo posiciona como uno de los mejores resultados en comparación con los otros modelos de Machine Learning. Este ajuste es especialmente relevante cuando las estadísticas de los conjuntos de entrenamiento y validación difieren significativamente, como en este caso.

A diferencia del modelo MLP, el modelo LSTM no tiene implementado el truncamiento de valores predichos mediante `np.clip()`, lo que significa que puede generar valores negativos en las predicciones. Esto es una diferencia clave con MLP, que ya tiene implementado este ajuste y, por lo tanto, evita predicciones incoherentes. Aunque LSTM mostró un buen ajuste con los datos reales, en escenarios donde es importante evitar predicciones fuera del rango esperado, sería recomendable implementar un ajuste como el truncamiento para asegurar la consistencia de los resultados entre los distintos modelos.

Discusión de aplicación de modelos en la industria

Tal como se comentó en la sección de antecedentes, muchos de los sistemas de control de la molienda SAG quedan sujetos a la experiencia de quien los opera, por lo que son susceptibles a eventos y desviaciones difíciles de prever para quien no los haya vivido antes.

De esta forma, se evalúan los beneficios de contar con un sistema de predicción de consumo de potencia confiable y su aplicación en sistemas de control avanzado, donde al contar con este valor de forma anticipada, se puede generar un sistema de control tal que al utilizar esta información pueda variar los límites de operación de forma automática, y así mantener estable el sistema en todo momento.

Una predicción del consumo con alta precisión en diversas ventanas temporales a tiempo real favorecería la detección temprana de eventos, permitiendo tomar medidas preventivas en lugar de reactivas. Así mismo, se podrían tomar acciones preventivas al encontrarse con perturbaciones del sistema, pudiendo anticipar y controlar perturbaciones en el sistema, el cómo afectarán estas variables al consumo de potencia y las acciones requeridas para controlar esta situación.

Para llegar a esto, se debe seguir trabajando en alcanzar una alta confiabilidad en la predicción de consumo de potencia a tiempo real, disminuyendo el error obtenido (rRMSE) y aumentando la correlación entre datos reales y predicción (CC), y generando un modelo que pueda ser utilizado para la toma de decisiones críticas dentro de la operación.

Junto con la predicción de eventos de sobrecarga, se puede evaluar la aplicación de estos modelos para la identificación de nivel de llenado de bolas, identificación de eventos de quiebre de revestimientos, optimización de tratamiento vs consumo de potencia, mejoramiento de parámetros de eficiencia energética, distribuciones de granulometría salientes, entre otros.

6. CONCLUSIONES

Este trabajo aplica distintos métodos de Machine Learning y Deep Learning para la predicción de consumo de potencia en la molienda SAG a tiempo real. Se aplica la metodología propuesta por Ávalos et al. (2020) para evaluar el consumo de potencia en molienda SAG con datos de la industria, comprobando que es exportable a otros conjuntos de datos de procesamiento de minerales con distinta temporalidad y variables, obteniendo resultados que mantienen las tendencias observadas en la bibliografía.

Es crucial contar con sensores confiables en los circuitos de molienda para asegurar la calidad de los datos y reducir pérdidas de información por fallas de medición o conexión.

Inicialmente, para caso 1, los mejores resultados se obtuvieron con Deep Learning, destacándose el modelo LSTM con 24 nodos, que alcanzó un error de 9.45% y una correlación de 0.83. Sin embargo, con la inclusión de variables operacionales, el mejor método de predicción pasó a ser MLP con 32 nodos, el cual incluyó las variables base FT, SSp y BPr, además del tonelaje de revestimiento FTR y la granulometría PS. Este modelo logró un error mínimo de 9.22% y una correlación de 0.83.

En los modelamientos mediante Machine Learning, el mejor desempeño se obtuvo con el modelo PR, utilizando un polinomio de grado 1 y las variables base, el tonelaje de campaña y la granulometría, con un rRMSE de 9.71% y un coeficiente de correlación (CC) de 0.82.

Se debe considerar que en código de modelo MLP, se utilizó `np.clip()` para truncar predicciones negativas, lo que afectó los valores de CC y rRMSE. Este ajuste no se aplicó en otros modelos, lo que generó inconsistencias en la comparación de resultados. Para una evaluación uniforme, se sugiere aplicar este ajuste en todos los modelos o en ninguno.

Aunque el error rRMSE se redujo levemente al agregar con nuevas variables, menos del 1%, se mantiene un alto grado de certeza al usar solo las variables base, lo que sugiere que podría realizarse el modelamiento solo con estas variables y mantener un resultado aceptable.

Los errores rRMSE fueron muy sensibles a la desviación estándar de los conjuntos de entrenamiento. El Set 7, con una desviación estándar de 3890 kWh, mostró un error de casi el doble comparado con el Set 3, con 2003 kWh.

Las diferencias estadísticas entre los conjuntos de entrenamiento y validación afectan los resultados al desnormalizar usando las estadísticas de entrenamiento, lo cual es inevitable. Esto puede generar valores incoherentes, como potencias negativas, afectando directamente a CC y rRMSE. Para mitigar este problema, es esencial mejorar la representatividad del conjunto de entrenamiento, abarcando periodos operacionales más diversos. Esta situación se da en particular en el entrenamiento utilizando sets de datos por período debido a su corta duración respecto a datos globales.

La implementación de un sistema de predicción confiable del consumo de potencia puede mejorar significativamente los sistemas de control avanzados en la industria minera. La capacidad de predecir el consumo energético en tiempo real permite la detección temprana de eventos y la toma de medidas preventivas, mejorando la estabilidad y eficiencia operativa.

Este trabajo no solo valida el uso de modelos de Machine Learning y Deep Learning en la predicción del consumo energético en molienda SAG, sino que también expande el conocimiento existente al incorporar nuevas variables y evaluar diferentes configuraciones y soportes temporales. Esto proporciona una base sólida para futuras investigaciones y aplicaciones en la industria minera, destacando la importancia de la adaptabilidad y precisión en los sistemas de predicción energética.

Finalmente, la metodología expuesta tiene el potencial de ser extendida a otros contextos mineros con diferentes variables operacionales y temporales, junto con ser evaluada para la aplicación dentro de sistemas de control automático.

Futuras Líneas de Investigación

Se recomienda continuar trabajando en la mejora de la confiabilidad de los modelos predictivos, enfocándose en la reducción del error rRMSE y el aumento de la correlación entre los datos reales y las predicciones.

Es clave explorar la posibilidad de entrenar los modelos con más datos que representen mejor los periodos de la base de validación, evitando diferencias entre las estadísticas de entrenamiento y validación. Esto puede implicar incluir ventanas temporales más amplias o datos más alineados entre los conjuntos.

Se debe explorar la combinación de Machine Learning y Deep Learning con modelos físicos tradicionales para mejorar la precisión y robustez de las predicciones. El desarrollo de algoritmos de optimización en tiempo real basados en modelos predictivos permitirá ajustar automáticamente los parámetros operacionales del molino, mejorando la eficiencia energética y reduciendo costos.

Modelar distintos escenarios operacionales y evaluar su impacto en el consumo energético y la eficiencia del proceso proporcionará una herramienta de apoyo para la toma de decisiones estratégicas. Futuras investigaciones podrían centrarse en cómo la optimización del consumo energético puede reducir la huella de carbono de las operaciones mineras, alineándose con los objetivos de sostenibilidad de la industria.

Implementando estas líneas de investigación, se avanzará en la precisión y aplicabilidad de los modelos predictivos, contribuyendo al desarrollo de operaciones mineras más eficientes y sostenibles

7. BIBLIOGRAFÍA

- [1] Avalos, S., Kracht, W. and Ortiz, J.M., 2020. Machine learning and deep learning methods in mining operations: A data-driven SAG mill energy consumption prediction application. *Mining, Metallurgy & Exploration*, 37(4), pp.1197-1212.
- [2] Bond, F.C., 1961. Crushing and grinding calculations *Brit. Chemical Engineering*, (8).
- [3] Hogg, R., 1972. Power relationships for tumbling mills. *AIME Trans.*, 252, pp.418-423.
- [4] Austin, L.G. and Luckie, P.T., 1972. Methods for determination of breakage distribution parameters. *Powder Technology*, 5(4), pp.215-222.
- [5] Morrell, S., 1996. Power draw of wet tumbling mills and its relationship to charge dynamics - Part 1: a continuum approach to mathematical modelling of mill power draw. *Trans Inst Min Metall, Section C Vol 105*, C43-53.
- [6] Morrell, S., 2009. Predicting the overall specific energy requirement of crushing, high pressure grinding roll and tumbling mill circuits
- [7] Casali, A. (2015). Modelling SAG milling power and specific energy consumption including the feed percentage of intermediate size particles.
- [8] Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H.A. and Kumar, V., 2018. Machine learning for the geosciences: Challenges and opportunities. *IEEE Transactions on Knowledge and Data Engineering*, 31(8), pp.1544-1554.
- [9] Bergen, K.J., Johnson, P.A., Maarten, V. and Beroza, G.C., 2019. Machine learning for data-driven discovery in solid Earth geoscience. *Science*, 363(6433).
- [10] Moraga, C., Kracht, W., & Ortíz, J. M. (2022). Process simulation to determine blending and residence time distribution in mineral processing plants. *Minerals Engineering*, 187, 107807.
- [11] Lishchuk, V. V., Lund, C., & Ghorbani, Y. (2019). Evaluation and comparison of different machine-learning methods to integrate sparse process data into a spatial model in geometallurgy. *Minerals Engineering*, 134, 156-165.
- [12] Mitchell, T., 1997. *Machine learning*
- [13] Ramezani, M., & Tafazoli, S. (2021). Using Artificial intelligence in mining excavators: Automating routine operational decisions. *IEEE Industrial Electronics Magazine*, 15(1), 6-11. <https://doi.org/10.1109/mie.2020.2964053>
- [14] Noriega, R., & Pourrahimian, Y. (2022). A Systematic review of artificial intelligence and data-driven approaches in strategic open-pit mine planning. *Resources Policy*, 77, 102727. <https://doi.org/10.1016/j.resourpol.2022.102727>
- [15] Wang, S., Zhang, Z., & Wang, C. (2023). Prediction of stability coefficient of open-pit mine slope based on artificial intelligence deep learning algorithm. *Scientific Reports*, 13(12017),1-10.

- [16] Choi, Y., Nguyen, H., & Bui, X. (2022). Optimization of haulage-truck system performance for ore production in open-pit mines using big data and machine learning-based methods. *Resources Policy*, 75, 102522.
- [17] Li, S., Sari, Y. A., & Kumral, M. (2020). Optimization of mining–mineral processing integration using unsupervised machine learning algorithms. *Natural Resources Research*, 29(5), 3035-3052.
- [18] Saldaña, M., Gálvez, E., Navarra, A., Toro, N., & Cisternas, L. A. (2023). Optimization of the SAG grinding process using statistical analysis and machine learning: A case study of the Chilean copper mining industry. *Materials*, 16(3220), 1-23.
- [19] McCoy, J. T., & Auret, L. (2019). Machine learning applications in minerals processing: A review. *Minerals Engineering*, 132, 95-109.
- [20] Estay, H., Ortiz, J. M., & García, C. (2023). On the Challenges of Applying Machine Learning in Mineral Processing and Extractive Metallurgy. *Minerals*, 13(2), 182.

ANEXO

Material complementario

Se realiza respaldo de códigos utilizados en repositorio GitHub para futuros desarrollos.

*[https://github.com/FC – Fer/Data – Driven – SAG – Mill – Energy – Consumption
– Prediction/tree/main](https://github.com/FC-Fer/Data-Driven-SAG-Mill-Energy-Consumption-Prediction/tree/main)*