

UNIVERSIDAD DE CONCEPCIÓN
Facultad de Ingeniería
Departamento de Ingeniería
Metalúrgica

Profesor Patrocinante
Dr. Fernando Betancourt Cerda

Ingeniero Supervisor
Nelson Mora Navarrete

DISEÑO DE MODELO PREDICTIVO MULTIVARIABLE PARA MOLIENDA EN COMPAÑÍA MINERA DOÑA INÉS DE COLLAHUASI



FRANCISCA VALENTINA CONTRERAS ACCHIARDO

Informe de Memoria de Título
para optar al Título de

Ingeniero Civil Metalúrgico

Agosto 2018

Agradecimientos

Agradezco a mis padres M^a Eugenia y Erwin por su apoyo y amor incondicional, por velar cada día por mí y por estar siempre presentes en cada etapa de mi vida, los amo incondicionalmente.

A mi abuelo Alfredo, por entregarme tanto, el amor a los libros y a los animales, enseñarme a vivir la vida valorando cada oportunidad y por siempre tener un recuerdo cariñoso en los momentos difíciles.

A Felipe, por ser un pilar, por aguantarme, calmarme y ayudarme en los momentos de estrés, por quererme como soy.

A mis amigas de la vida Natalia, M^a José y Laura, por esos cigarritos de 20 minutos, por escuchar, entender y reír de nuestros fracasos.

A mis compañeros y amigos de la U, Pía, Carla, Nelson, Fabián y Danor, gracias por todas las experiencias compartidas, los apretes, vigiliias y carretes, hicieron de la universidad un lugar especial y de lindos recuerdos.

A mi amiga que me dejó Iquique, Valentina, por hacernos compañía mutuamente y estar para mí cuando no conocía a nadie en la ciudad.

A mi profesor guía Fernando Betancourt, por creer en mis capacidades y acompañarme durante el desarrollo de esta memoria.

A CMDIC, Gerencia de Geometalurgia, por darme la oportunidad de desarrollarme como profesional junto a ustedes, a mi tutor Nelson Mora por enseñarme con paciencia y manzanitas y a José Ignacio Navarro por ser un amigo y profesor cuando lo necesitaba.

Resumen

El presente trabajo se desarrolló en las dependencias de la gerencia de geometalurgia de la Compañía Minera Doña Inés de Collahuasi, con el objetivo de desarrollar un modelo que fuera capaz de predecir el comportamiento no lineal de la potencia de uno de los molinos SAG de CMDIC, teniendo como entradas del modelo las variables mineralógicas (durezas y granulometría de alimentación) junto con las variables manipulables por el operador (%sólidos, velocidad, nivel de bolas y flujo de alimentación) y como salida única la potencia.

La idea surge debido a que la planta no cuenta con una estimación confiable de la potencia en los molinos SAG, lo que conlleva a una variabilidad en la operación, error en las planificaciones y estimaciones de tratamiento.

Se proponen tres tipos de modelos predictivos a comparar: Modelo de Redes Neuronales y Modelo de Bosques Aleatorios, los cuales fueron creados y validados utilizando el software estadístico R, junto con el modelo de potencia establecido por Austin. Para la realización de los modelos se utilizó data real recopilada de planta, estos datos fueron obtenidos desde mediciones en terreno que fueron reportados al servidor Pi ProcessBook. Previamente al entrenamiento de los modelos, la data cruda se sometió a un tratamiento en donde se eliminaron los valores atípicos u outliers.

El criterio de comparación de los modelos fue la Raíz del Error Cuadrático Medio (Root Mean Square Error, RMSE) y el Error Porcentual Absoluto Medio (Mean Absolute Percentage Error, MAPE).

De los resultados se concluye que el método de Random Forest con 150 árboles de decisión tiene un mejor desempeño que el modelo de Redes Neuronales con 2 neuronas en la capa oculta. Mientras los bosques tienen un error de 0,76%; el de las redes es 1,40 %. Por otro lado el modelo de Austin posee un error de 3,04%.

Finalmente, como la potencia depende principalmente del nivel de llenado total del molino, se propone como trabajo futuro desarrollar un modelo predictivo de nivel de llenado total, lo que contribuye de mejor manera con el modelamiento de la potencia del molino.

Abstract

The present work was developed in the dependencies of the management of geometallurgy, of the Compañía Minera Doña Inés de Collahuasi with the objective of developing a model that was able to predict the non-linear behavior of the power of one of the SAG mills of CMDIC, having as inputs of the model the mineralogical variables (hardness and feed particle size) together with the variables manipulated by the operator (% solids, speed, ball level and throughput) and as a single output the power.

The idea arises because the plant does not have a reliable estimate of the power in the SAG mills, which leads to a variability in the operation, error in the planning and treatment estimates.

Three types of predictive models to be compared are proposed: Model of Neural Networks and Model of Random Forests, which were created and validated using the statistical software R, together with the established power model of Austin. For the realization of the models real data collected from the plant was used, these data were obtained from field measurements that are reported to Pi ProcessBook server. Prior to the training of the models, the raw data underwent a treatment in which atypical values were eliminated.

The criterion for comparing the models was the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE).

From the results it is concluded that the Random Forest method with 150 decision trees has a better performance than the Neural Networks model with 2 neurons in the hidden layer. While the forests have an error of 0,76%; that of the networks is 1,40%. On the other hand the Austin model has an error of 3,04%.

Finally, as the power mainly depends on the level of total filling of the mill, it is proposed as future work to develop a predictive model of total filling level, which contributes in a better way with the modeling of the power of the mill.

Nomenclatura

- **Acrónimos**

BChy:	Batería de hidrociclones.
CMDIC:	Compañía Minera Doña Inés de Collahuasi.
CEE:	Consumo Específico de Energía.
HC:	Hidrociclones.
MAPE:	Mean Absolute Percentage Error o Error Porcentual Absoluto Medio.
MISO:	Multiple inputs single output.
MB:	Molinos de Bolas.
RF:	Random Forest.
RMSE:	Root Mean Square Error o Raíz del Error Cuadrático Medio.
RNA:	Redes neuronales artificiales.
SAG:	Semi autógeno.

- **Unidades de Medida**

km:	kilómetros
kW:	kilo watts
kWh/m ³ :	kilo watts hora por metro cúbico
m:	metros
m ³ /h:	metros cúbicos por hora
m.s.n.m:	Metros sobre el nivel del mar
pulg:	pulgadas
rpm:	revoluciones por minuto
ton:	toneladas
tph:	toneladas por hora
t/m ³ :	toneladas por metro cúbico



- **Nomenclatura Modelo de Austin:**

P_N : Potencia neta en el piñón de un Molino cilíndrico o mill shell

K: Constante de sintonización

D: Diámetro del molino dentro de los liners

L: Largo efectivo

A: Constante de sintonización

ε_b : Porosidad de la cama de bolas y rocas

ρ_i : Densidad del componente i

J_i : Fracción de llenado del componente i

w_c : Fracción en peso de sólido de la pulpa

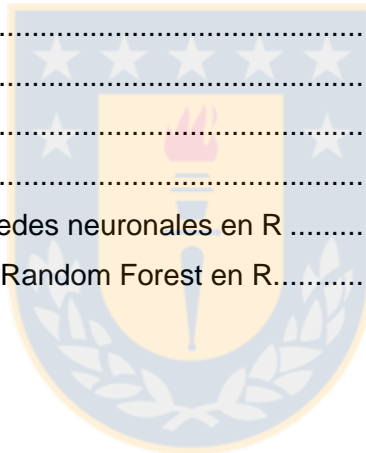
N_c : Fracción de la velocidad crítica de giro



Índice

Resumen.....	i
Abstract.....	ii
Nomenclatura.....	iii
Índice de Tablas.....	vii
Índice de Figuras.....	viii
1 Introducción.....	1
1.1 Objetivos.....	2
1.1.1 Objetivo General.....	2
1.1.2 Objetivos Específicos.....	2
1.2 Alcances y Limitaciones.....	3
1.2.1 De los datos utilizados para la generación de los modelos.....	3
1.2.2 De la generación de los modelos.....	3
2 Antecedentes de la Empresa.....	4
2.1 Descripción de la Empresa.....	4
2.2 Descripción de la Operación.....	5
2.3 Etapas de Molienda y Clasificación.....	6
2.3.1 Línea 3 de Molienda en CMDIC.....	7
3 Estado de Arte.....	8
3.1 Molienda SAG.....	8
3.2 Potencia en los molinos SAG.....	9
3.3 Factores que afectan la operación de un molino SAG.....	11
3.3.1 Flujo de Alimentación.....	11
3.3.2 Distribución Granulométrica en la alimentación.....	13
3.3.3 Dureza del mineral.....	14
3.3.4 Carga de Bolas.....	14
3.3.5 Densidad y viscosidad de la pulpa.....	15
3.4 Modelos de potencia.....	16
3.4.1 Modelo de Austin.....	16
3.5 Modelación de Sistemas.....	17
3.5.1 Clases de Modelos.....	17
3.5.2 Redes Neuronales Artificiales.....	19
3.5.3 Bosques Aleatorios.....	22
3.6 Validación de los modelos.....	25
3.6.1 Validación Cruzada (Cross Validation).....	26

3.7	Eliminación de Outliers o Valores Atípicos.....	27
3.7.1	Test de Tukey	27
4	Desarrollo Experimental	28
4.1	VARIABLES involucradas en el modelo	28
4.2	Filtro de data cruda	28
4.3	Búsqueda de Nivel de Llenado (J) para Modelo de Austin	31
4.4	Búsqueda de la distribución de J en Modelo de Austin.....	33
5	Resultados	34
5.1	Eliminación de Outliers.....	34
5.2	Generación de modelos con Redes Neuronales	37
5.3	Generación de modelo utilizando Random Forest.....	42
5.4	Modelo de Austin	44
5.5	Desempeño de Modelos	45
6	Conclusiones.....	47
7	Bibliografía	49
8	ANEXOS	51
8.1	Programa R.....	51
8.2	Generación modelo de redes neuronales en R	51
8.3	Generación de Modelos Random Forest en R.....	55



Índice de Tablas

Tabla 1: Características principales del molino SAG 1011 de CMDIC	7
Tabla 2: Parámetros utilizados para el modelo de Austin	31
Tabla 3: Resumen de las variables numéricas en el conjunto de datos disponibles	36
Tabla 4: Parámetros a considerar para la generación de los modelos RNA's	37
Tabla 5: Número de neuronas versus errores RSME de entrenamiento, validación y validación cruzada con $k = 10$ iteraciones	37
Tabla 6: Errores de Red Neuronal con 2 neuronas.....	39
Tabla 7: Pesos correspondientes a cada conexión entre las entradas y la neurona 1 de la capa oculta	41
Tabla 8: Pesos correspondientes a cada conexión entre las entradas y la neurona 2 de la capa oculta	41
Tabla 9: Valores de bias de las neuronas en la capa oculta y capa de salida de la mejor red neuronal	41
Tabla 10: Pesos que conectan las neuronas de la capa oculta con la capa de salida	41
Tabla 11: Parámetros a considerar para la generación de los modelos random forest.....	42
Tabla 12: Errores correspondientes a Random Forest con 150 árboles de decisión	43
Tabla 13: Errores asociados al modelo de Austin con J constante	44



Índice de Figuras

Figura 1: Ubicación Compañía Minera Doña Inés de Collahuasi	4
Figura 2: Diagrama de Proceso planta sulfuros CMDIC	5
Figura 3: Circuito de Molienda y Clasificación en CMDIC	6
Figura 4: Esquemización Línea 3 de molienda en CMDIC	7
Figura 5: Esquema de un molino SAG	8
Figura 6: Representación idealizada de la carga de un molino en operación [3]	9
Figura 7: Demanda de potencia del molino semiautógeno en función del nivel de llenado [3]	10
Figura 8: Relación entre el flujo de alimentación y el volumen de la carga en un molino [3]	11
Figura 9: Efecto del flujo de alimentación en la potencia neta consumida por el molino [3]	12
Figura 10: Efecto de la carga fresca (rocas irregulares) y vaciado de una sobrecarga (rocas redondeadas) sobre la demanda de potencia del molino [3].....	12
Figura 11: Arquitectura de una neurona artificial	19
Figura 12: RNA multilayer feedforward	20
Figura 13: Representación de un árbol de decisión	22
Figura 14: Esquemización de Random Forest.....	24
Figura 15: Representación de un gráfico de caja.....	27
Figura 16: Esquema de variables de entrada y salida en modelo MISO	28
Figura 17: Gráfica de caja para 'Potencia' sin eliminación de outliers.....	29
Figura 18: Gráfica de caja para 'Tratamiento' sin eliminación de outliers	29
Figura 19: Gráfica de caja para 'Velocidad' sin eliminación de outliers.....	29
Figura 20: Histograma para 'Potencia' sin eliminación de outliers	29
Figura 21: Histograma para 'Tratamiento' sin eliminación de outliers	29
Figura 22: Histograma para 'Velocidad' sin eliminación de outliers	29
Figura 23: Grafica de caja para 'Porcentaje de Sólidos' sin eliminación de outliers	30
Figura 24: Grafica de caja para 'Tamaño -2" ' sin eliminación de outliers	30
Figura 25: Histograma para 'Porcentaje de Sólidos' sin eliminación de outliers	30
Figura 26: Histograma para 'Porcentaje de Sólidos' sin eliminación de outliers	30
Figura 27: Histograma para 'Tamaño -2" ' sin eliminación de outliers	30
Figura 28: Histograma para 'Nivel de Bolas' sin eliminación de outliers.....	30
Figura 29: Gráfica de caja para 'Dureza DWI' sin eliminación de outliers	31
Figura 30: Histograma para 'Dureza DWI ' sin eliminación de outliers	31
Figura 31: Extracto de Excel utilizado para el cálculo de nivel de llenado total	32
Figura 32: Solver utilizado para encontrar el valor de J que minimiza el error.	32
Figura 33: Extracto de excel utilizado para encontrar la distribución de nivel de llenado total	33
Figura 34: Macros utilizadas para solución de solver	33
Figura 35: Gráfica de caja para 'Potencia' sin presencia de outliers.....	34
Figura 36: Gráfica de caja para 'Tratamiento' sin presencia de outliers.....	34
Figura 37: Histograma para 'Potencia' sin presencia de outliers.....	34
Figura 38: Histograma para 'Tratamiento' sin presencia de outliers.....	34
Figura 39: Gráfica de caja para 'Velocidad' sin presencia de outliers	35
Figura 40: Gráfica de caja para 'Porcentaje de Sólidos' sin presencia de outliers	35
Figura 41: Gráfica de caja para 'Tamaño -2" ' sin presencia de outliers	35
Figura 42: Histograma para 'Velocidad' sin presencia de outliers.....	35
Figura 43: Histograma para 'Porcentaje de Sólidos' sin presencia de outliers	35
Figura 44: Histograma para 'Tamaño -2" ' sin presencia de outliers	35
Figura 45: Gráfica de caja para 'Nivel de Bolas' sin presencia de outliers	36
Figura 46: Gráfica de caja para ' Dureza DWI' sin presencia de outliers.....	36

Figura 47: Histograma para 'Nivel de Bolas' sin presencia de outliers.....	36
Figura 48: Histograma para 'Dureza DWI' sin presencia de outliers	36
Figura 49: Raíz de error cuadrático medio de entrenamiento y validación cambiando el número de neuronas en la capa oculta	38
Figura 50: Raíz del error cuadrático medio de validación cruzada cambiando el número de neuronas en la capa oculta	39
Figura 51: Representación de la red neural con 2 neuronas en la capa oculta.....	40
Figura 52: Error cuadrático medio (con datos escalados) versus número de árboles	42
Figura 53: Importancia de las variables en el modelo RF	43
Figura 54: Distribución del nivel de llenado total del molino.....	44
Figura 55: Histograma de error para modelo redes neuronales.....	45
Figura 56: Histograma de error para modelo random forest	46
Figura 57: Histograma de error para modelo de Austin	46



1 Introducción

Uno de los insumos más utilizados en la minería y de mayor costo en la actualidad es la energía, siendo los equipos de molienda como el molino SAG, los molinos de barras y molinos de bolas aquellas máquinas que consumen la mayor cantidad de energía y de forma ineficiente; debido a sus grandes dimensiones y a los medios de molienda que utilizan.

Por este motivo, las empresas buscan optimizar el gasto energético producido por estos equipos ya que un bajo porcentaje de la energía que consumen es realmente utilizada para moler el mineral; mientras que el resto se utiliza para mover el molino y los medios de molienda, además de algunas pérdidas energéticas.

En planta se define el consumo específico de energía como la relación entre los kilowatt por hora que se consumen al moler una tonelada de un cierto mineral. Esta variable así definida C.E.E., se utiliza para evaluar el rendimiento energético de la planta.

Por lo tanto, suena coherente realizar estrategias operacionales para poder predecir la potencia frente a diferentes escenarios operacionales, ya sea en variabilidad de mineral como variables operacionales con el fin de maximizar el tratamiento del molino. Esto último para responder a la disminución de la ley del mineral y al continuo aumento en los costos de producción.

En la actualidad CMDIC no posee modelos de potencia asociados a los molinos SAG, por lo que se propone crear un modelo multivariable para el molino SAG 1011 de la línea de molienda 3 de la compañía (el cual trata alrededor del 60% del mineral de la compañía), utilizando no sólo variables manipulables por el operador, sino también considerar las variables mineralógicas pertenecientes a las rocas provenientes del stockpile.

1.1 Objetivos

1.1.1 Objetivo General

Obtener un modelo predictivo multivariable que permita determinar el comportamiento futuro de la potencia de los molinos SAG de CMDIC, frente a variaciones mineralógicas y variables manipulables.

1.1.2 Objetivos Específicos

- Adquirir y filtrar los datos para la generación de los modelos.
- Elaborar un código programable que permita obtener de manera iterativa los distintos modelos de Bosques Aleatorios y de Redes Neuronales que resulten al hacer variar los parámetros necesarios para su generación.
- Comparar los modelos desarrollados con modelo estándar de la literatura.



1.2 Alcances y Limitaciones

1.2.1 De los datos utilizados para la generación de los modelos.

Los datos utilizados en esta tesis son reales de la planta. Su recolección se hizo a través del servidor PI ProcessBook que CMDIC pone a disposición de los trabajadores para acceder a todos los datos de planta de manera práctica y sencilla, permitiendo almacenarlos en una planilla Excel.

Para el desarrollo de esta tesis se escogieron sólo los datos del molino SAG 1011 que fueran representativos de una semana de operación normal. Es decir, que no se presenten condiciones no deseadas en el proceso, como detenciones operacionales. La idea es que se consideren sólo los puntos operacionales que permitan predecir el comportamiento del molino sin imprevistos que dificulten la generalización del modelo.

No se consideran los datos de los molinos SAG 1 y 2. A pesar de que en teoría el material entrante posee las mismas características mineralógicas para los tres molinos SAG, éstos se encuentran dimensionados de manera diferente y poseen un circuito distinto, por lo que es necesario obtener un modelo para cada molino que quedará como trabajo futuro.

1.2.2 De la generación de los modelos.

Los códigos programables para la creación de los modelos de Redes Neuronales (sección 3.5.2) y Random Forest (sección 3.5.3) se desarrollaron con el software estadístico R [20], el cual proporciona un toolbox o librería con todas las funciones necesarias para la creación de ambos tipos de modelo.

Para que sea posible la comparación de los modelos Random Forest y Redes Neuronales, éstos se generaron utilizando los mismos datos en entrenamiento y testeo. El criterio de comparación del desempeño de los modelos es la raíz del error cuadrático medio (RMSE, Root Mean Square Error) utilizando un set de datos nuevo para seleccionar el mejor modelo.

En la práctica, se suelen usar redes neuronales con una o dos capas ocultas [5]. Por otro lado, no hay un criterio objetivo y universalmente aceptado acerca de la fórmula de selección óptima del número de neuronas ocultas. En [12] se sugiere que el número de neuronas ocultas en una red neuronal de tres capas debe ser de 75% del número de neuronas de entrada mientras que en [9] se propone duplicar el número de neuronas ocultas hasta que el rendimiento de la red comience a deteriorarse. Sin embargo, no se pueden dar reglas concretas y generales para determinar el número de neuronas o número de capas de una red para resolver un problema concreto, por lo que se debe ensayar [18].

2 Antecedentes de la Empresa

2.1 Descripción de la Empresa

Compañía Minera Doña Inés de Collahuasi se encuentra ubicada en la primera Región de Tarapacá, dividiendo sus operaciones en zona cordillerana entre los 4.400 y 4.900 msnm a unos 180 km al sureste de Iquique en la comuna de Pica y en la zona Puerto ubicado a 65 km al sur de Iquique en puerto Patache, como se observa en la Figura 1.

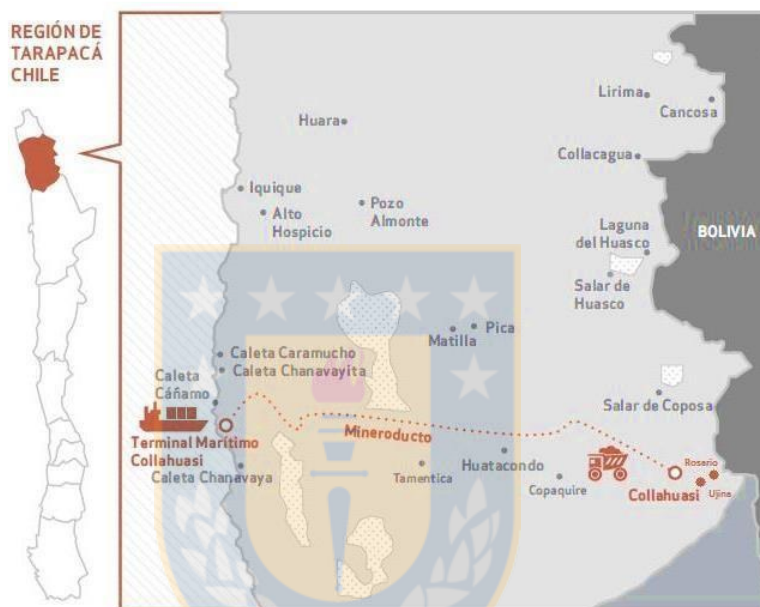


Figura 1: Ubicación Compañía Minera Doña Inés de Collahuasi

CMDIC inició sus operaciones en 1999, por medio de una sociedad contractual perteneciente a Anglo American y Glencore, cada una con un 44% de la propiedad, y el 12% restante pertenece a Japan Collahuasi Resources BV. Su principal negocio es la extracción y producción de concentrado de cobre y molibdeno extraídos desde sus principales yacimientos: Rosario y Ujina. Para el año 2017 CMDIC produjo un total de 520.704 toneladas de cobre fino representando un 9,5% de la producción total en Chile posicionándose en el tercer lugar de producción nacional después de CODELCO y Minera Escondida.

2.2 Descripción de la Operación

El mineral extraído desde el yacimiento Rosario y los diferentes stocks de la compañía, son enviados a la planta concentradora de sulfuros, donde se produce el concentrado de cobre y molibdeno, luego el concentrado es enviado a puerto Patache por medio del mineroducto, donde se realiza la flotación selectiva de molibdeno. Actualmente todo el mineral es tratado en la concentradora, ya que la línea de óxidos de lixiviación se encuentra paralizada desde el año 2015. En la Figura 2 se puede observar el diagrama de proceso de la planta de sulfuros de la compañía.

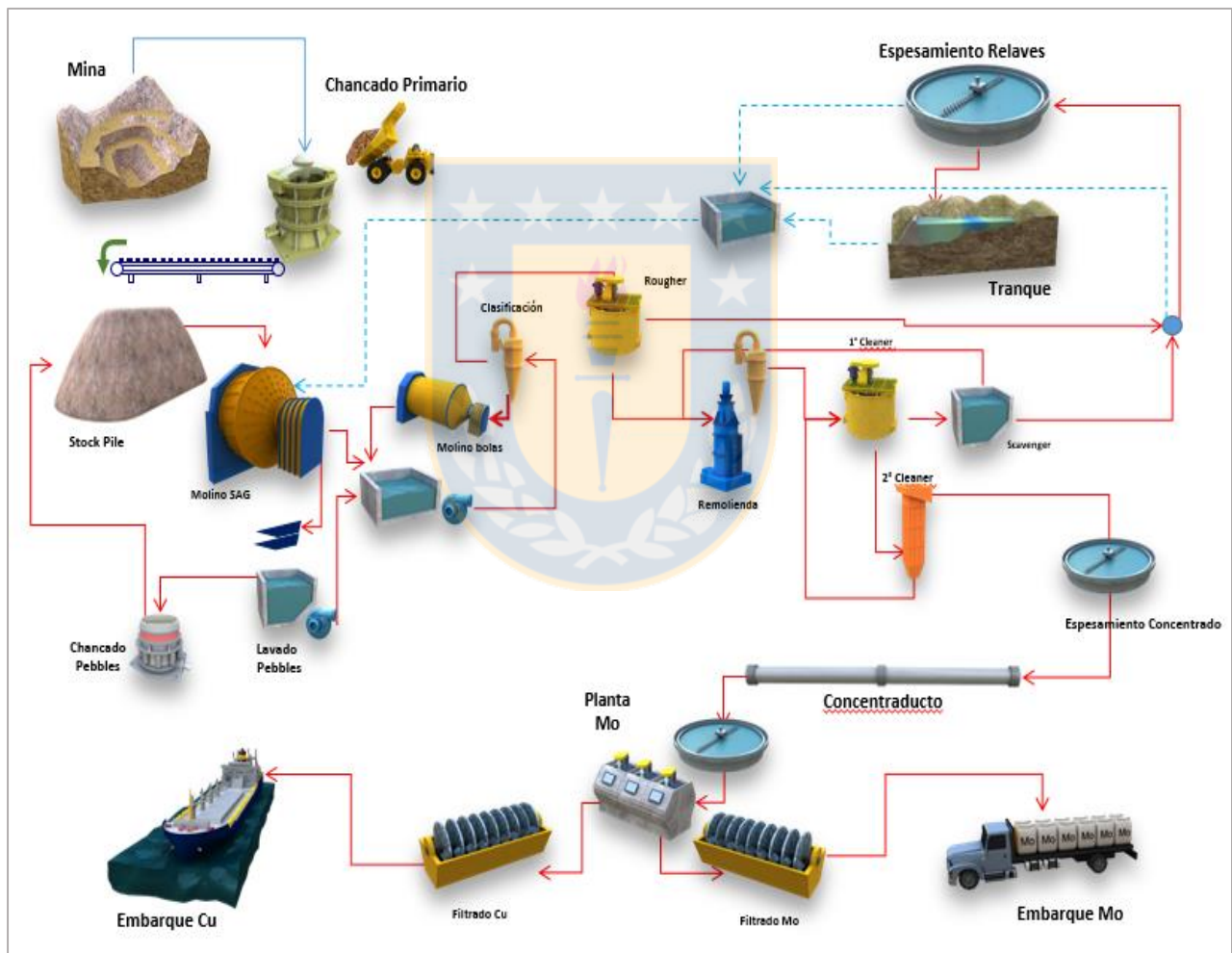


Figura 2: Diagrama de Proceso planta sulfuros CMDIC

2.3 Etapas de Molienda y Clasificación

La etapa comienza con el acopio de mineral grueso proveniente desde chancado primario. Con la ayuda de feeders y correas transportadoras el mineral es transportado a las 3 líneas de molienda. CMDIC cuenta con dos líneas de molienda gemelas (línea 1 y 2) que están compuestas por un molino SAG con capacidad aproximada de 1.500 TPH conectados en serie con un molino de bolas y una tercera línea de molienda que corresponde a la más nueva e importante de la compañía y se encuentra compuesta por un molino SAG con capacidad aproximada de 5.000 TPH conectado en serie con dos molinos de bolas (paralelos entre sí), en la actualidad se encuentra en estudio la incorporación de un quinto molino de bolas.

Cada una de las líneas recibe el mineral grueso desde el stockpile mezclado con la alimentación de las bolas de acero de 5,5" para luego en la boca del molino SAG ser mezclado con agua de proceso (75% agua de proceso y 25% agua fresca). La descarga de los molinos pasa a través de un trommel, que clasifica los tamaños intermedios o pebbles que puedan perjudicar el circuito enviándolos a la planta de chancado de pebbles la cual descarga al stockpile sector L1 y L2, mientras que la fracción fina de la pulpa es llevada a una cuba, mezclándose con la descarga de los molinos de bolas y agua de proceso.

Desde las cubas de alimentación, bombas extraen mezcla a un caudal de 2.500-4.800 [m³/h], cada una alimentando a una batería de 10 hidrociclones, donde 8 siempre se encuentran en uso dejando 2 en stand by por si alguno deja de funcionar. El overflow es dirigido a la etapa de flotación mientras el underflow alimenta a los molinos de bolas. El circuito se muestra en la Figura 3.

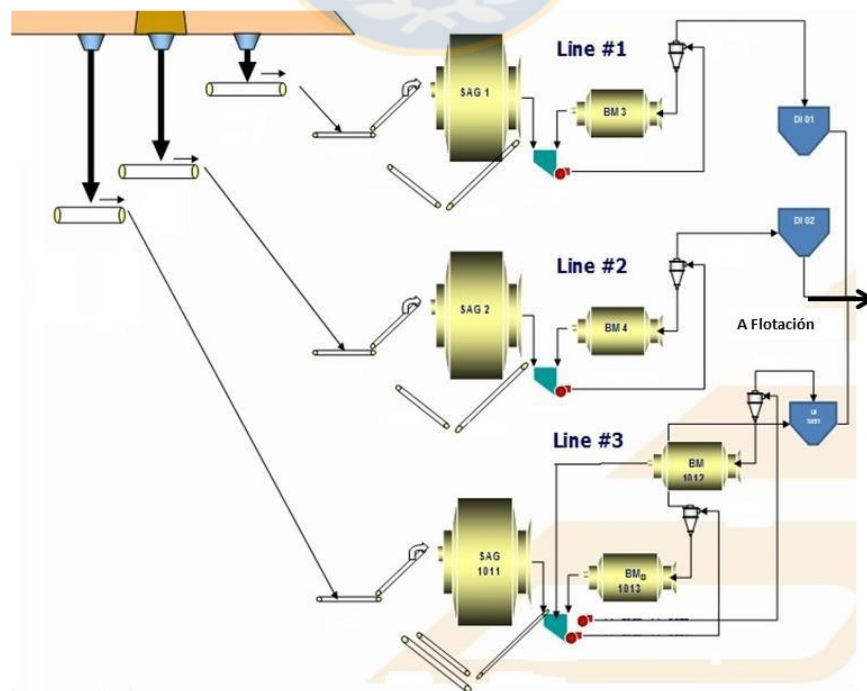


Figura 3: Circuito de Molienda y Clasificación en CMDIC

2.3.1 Línea 3 de Molienda en CMDIC

Dentro del circuito, esta memoria se enfocará en el molino SAG 1011 perteneciente a la línea 3, tomando datos reales de operación. A diferencia de las dos primeras líneas, la línea 3 posee un molino SAG con características diferentes las cuales quedan representadas en la Tabla 1:

Tabla 1: Características principales del molino SAG 1011 de CMDIC

Molino SAG1011		Unidad	Valor
Equipos	Marca		Svedala
	Diámetro	pies	40
	Largo	pies	22
	Potencia Nominal Motor	kW	23.000
	Velocidad crítica	%	74-80
	Tamaño Bolas	pulg	5,5
	Parrilla de descarga	pulg	2,5
	Malla	tipo	Trommel
	Slot Trommel	mm	25x50
KPI's	Rendimiento	TPH	5.000
	Tratamiento	TPD	120.000

Las diferencias operacionales en las líneas corresponden a la presencia de mayor cantidad de equipos secundarios, ya que para maximizar el rendimiento del molino SAG se cuenta con la ayuda de 2 molinos de bolas ubicados en paralelo entre sí (a diferencia de L1 y L2 que poseen solo uno), 4 baterías de HC y una eliminación total de los pebbles en el sistema (los pebbles generados serán enviados a chancado de pebbles y dirigidos al sector de stockpile que alimenta L1 y L2). En la Figura 4 queda representado de manera esquemática la operación de la L3:

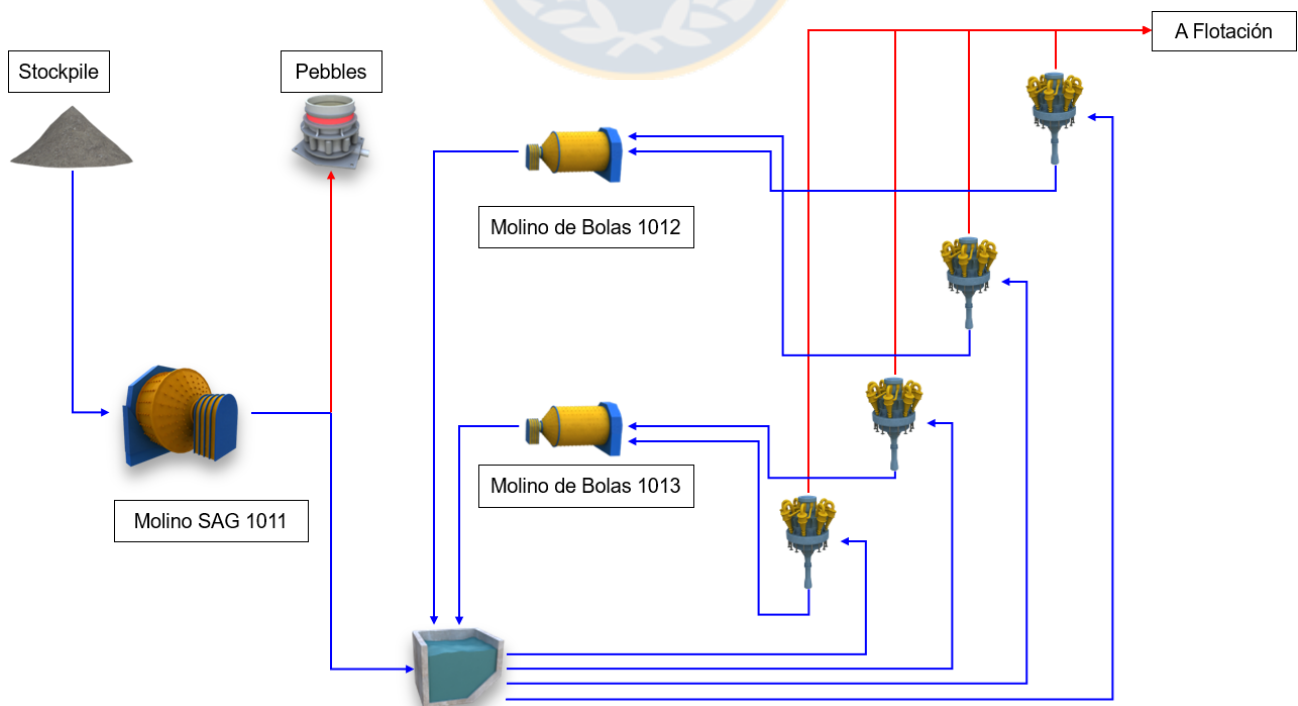


Figura 4: Esquematización Línea 3 de molienda en CMDIC

3 Estado de Arte

3.1 Molienda SAG

La molienda semi-autógena o molienda SAG es encargada de reducir el tamaño proveniente de la chancadora primaria. Esta se realiza en molinos rotatorios de gran diámetro y utiliza como medio de molienda la roca misma y una pequeña cantidad de bolas de acero, adicionalmente se agrega agua a la entrada del molino para usarla como medio de transporte para las partículas.

El mineral grueso es esencial para producir fractura de los tamaños más finos, la falta o disminución de estos puede provocar la acumulación de las partículas pequeñas y/o intermedias provocando así una disminución de la capacidad de molienda. El problema asociado a la falta de medios de molienda como medios de fractura se conoce como acumulación de tamaño crítico, donde el tamaño crítico corresponde al tamaño de mineral que es demasiado pequeño para ser medio de molienda y demasiado grande para ser fracturado por otras rocas, a este tamaño de mineral se le llama pebbles. Por ello la molienda SAG contiene bolas de acero, con el fin de ayudar a la eliminación del tamaño crítico.

Si describimos al molino SAG podemos decir que consiste en una cámara metálica cilíndrica giratoria en la cual la alimentación entra continuamente a través de un extremo y es descargada a través de otro extremo después de haber sido molido al tamaño deseado. El molino SAG está constituido por una carcasa que es protegida internamente por revestimientos de acero, una malla que permite a las partículas de mineral molido y agua pasar en forma de pulpa y un levitador de pulpa que corresponde a un mecanismo que como su nombre lo dice, levanta la pulpa que pasa a través de la malla para descargarla. En la Figura 5 queda esquematizada la estructura típica de un molino SAG.

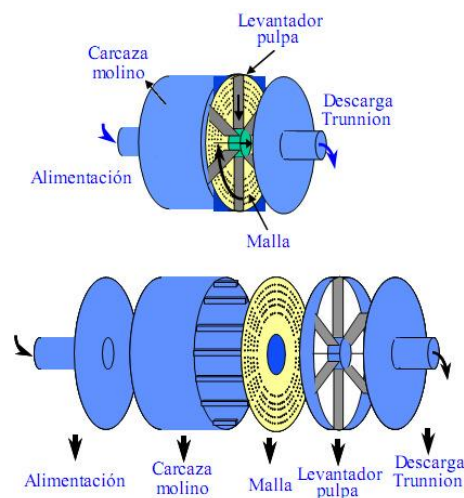


Figura 5: Esquema de un molino SAG

3.2 Potencia en los molinos SAG

La potencia P necesaria para rotar un molino es una de las variables operacionales de mayor importancia en la molienda semi-autógena. Se hace necesario entonces conocer de qué manera las variables operacionales afectan su valor. Para analizar esto haremos uso de la Figura 6. En ella se muestra un molino de diámetro interno D rotando a una velocidad N revoluciones por minuto (rpm) con una carga de mineral, bolas y pulpa de un peso M_t toneladas.

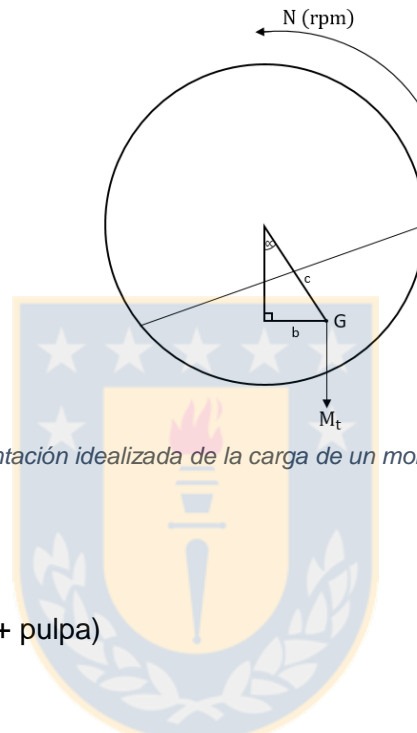


Figura 6: Representación idealizada de la carga de un molino en operación [3]

Donde,

N : Velocidad de rotación [rpm]

M_t : Peso total de la carga (bolas + pulpa)

G : Centro de masa

b : Distancia entre centro de masa G y el eje vertical de simetría del molino.

Para mantener el molino en rotación debemos ejercer un torque que es el proporcional al producto entre el peso M_t y la distancia o brazo b . Conociendo como varían estas cantidades con las condiciones de operación, podremos saber cómo se ve afectada la potencia. En un primer análisis se puede establecer lo siguiente:

1. El producto $M_t b$ entrega el torque que se necesita ejercer para mantener el molino en movimiento.
2. El brazo b aumenta con el ángulo de reposo. En consecuencia cualquier factor que afecte este ángulo afectará del mismo modo la potencia.
3. A medida que el nivel del molino aumenta, M_t se incrementa y b disminuye. Si el molino se encuentra vacío, el factor M_t es cero, y si está completamente lleno entonces el factor b es cero, es decir en ambos casos el torque $M_t b$ es cero. Por lo tanto, debe existir entre estos dos extremos un nivel de llenado del molino para el cual la potencia tiene un valor máximo como se muestra en la Figura 7.

4. Para un peso M_t constante, si la carga tiene una mayor densidad ocupará menos volumen y b se incrementará con lo cual la potencia se hace mayor.
5. Para el volumen de llenado constante, si la carga tiene una mayor densidad aumentará su peso M_t y la potencia aumentará.

Es importante notar que pequeñas variaciones en la capacidad de levantar su carga afectan considerablemente la potencia del molino. Por ejemplo si α aumenta de 40° a 45° la potencia aumentará en un 10% [3], si el resto de las condiciones permanecen constantes. Además los puntos 4 y 5, mencionados anteriormente muestran la importancia de la densidad de la carga para determinar la potencia demandada por el molino. También es importante mencionar que para una carga fija de bolas, la densidad varía con el nivel de llenado. Este efecto se hace sentir en la potencia y se muestra en la Figura 7:

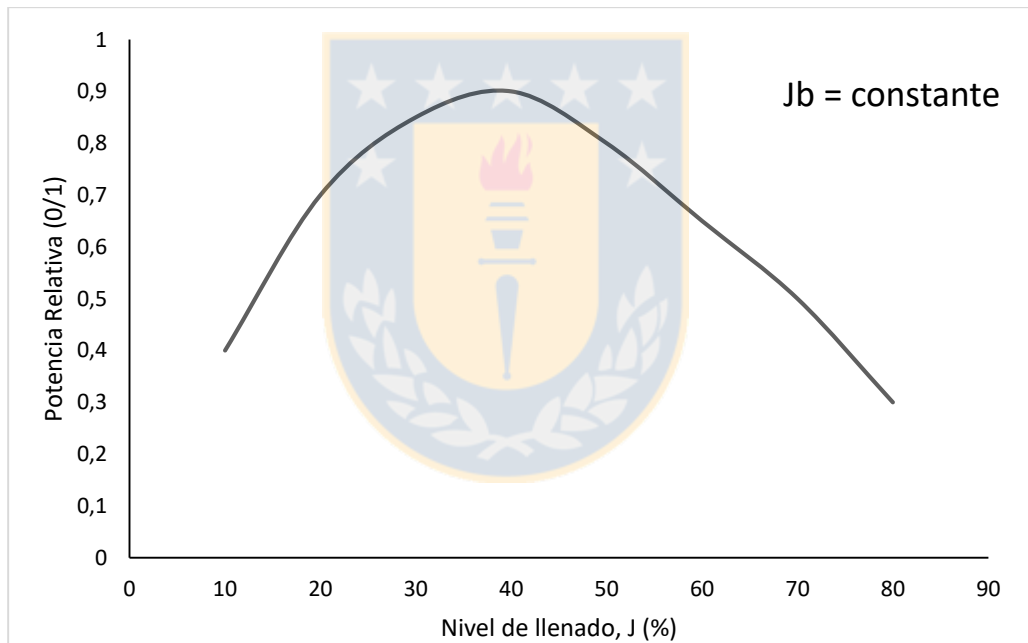


Figura 7: Demanda de potencia del molino semiautógeno en función del nivel de llenado [3]

3.3 Factores que afectan la operación de un molino SAG

3.3.1 Flujo de Alimentación

El grado de llenado de un molino SAG, depende del flujo de alimentación de mineral fresco con que opera [3]. Mientras mayor sea el flujo de alimentación, mayor será el volumen de la carga con que trabaja el molino. Esto se debe a que la molienda y descarga del mineral, son procesos cinéticos, en los cuales si las condiciones operacionales permanecen constantes, las masas de mineral molido y descargado por unidad de tiempo, son proporcionales a la masa presente en el molino. En consecuencia para igualar o balancear un aumento del flujo de entrada, la cantidad de mineral presente en el molino debe necesariamente aumentar. Esto ocurre así hasta un cierto valor del llenado por sobre el cual el proceso se revierte.

Debido a que el volumen de la carga está relacionado con el flujo de alimentación, como se indica en la Figura 8, en la práctica el nivel de la carga se controla ajustando el flujo de alimentación.

Además de la relación mencionada, el volumen de la carga tiene un efecto directo en la potencia, de tal manera que el flujo de alimentación y la potencia quedan relacionados como se indica en la Figura 9.

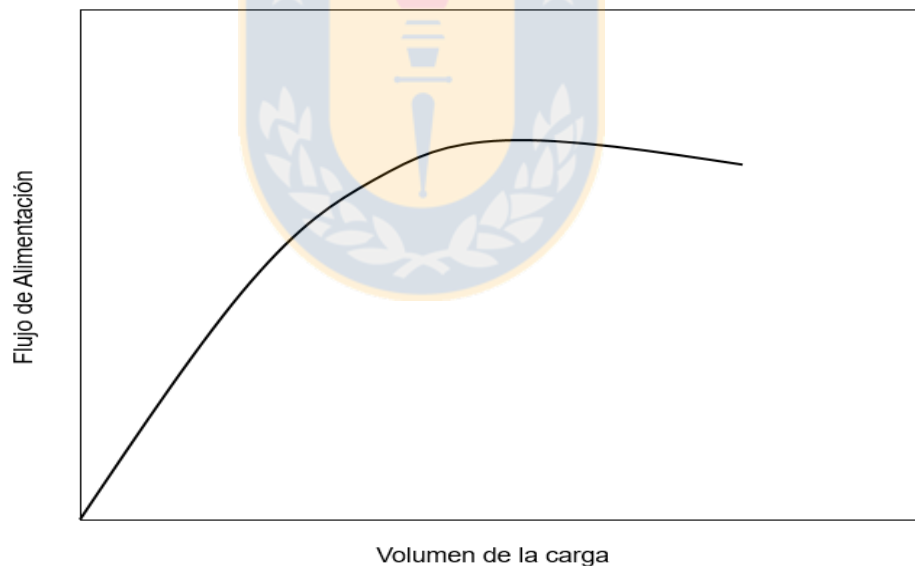


Figura 8: Relación entre el flujo de alimentación y el volumen de la carga en un molino [3]

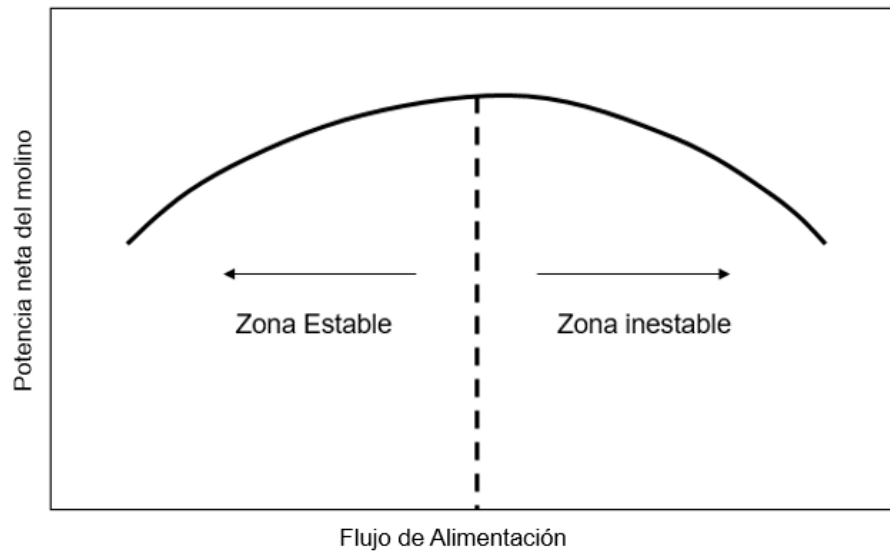


Figura 9: Efecto del flujo de alimentación en la potencia neta consumida por el molino [3]

En la figura anterior se observa que la potencia aumenta con el flujo, comenzando de un valor cero. A medida que el flujo de alimentación crece, la potencia consumida se incrementa hasta llegar a un valor máximo. Un flujo de alimentación aún mayor provocará una sobrecarga, la intensidad de la acción de molienda se reduce y la capacidad de tratamiento del molino disminuye. Frente a esta situación el operador parará la alimentación de sólidos y permitirá que se vacíe ("grind out"). Luego reanudará la alimentación a una tasa más baja que permita una operación estable nuevamente.

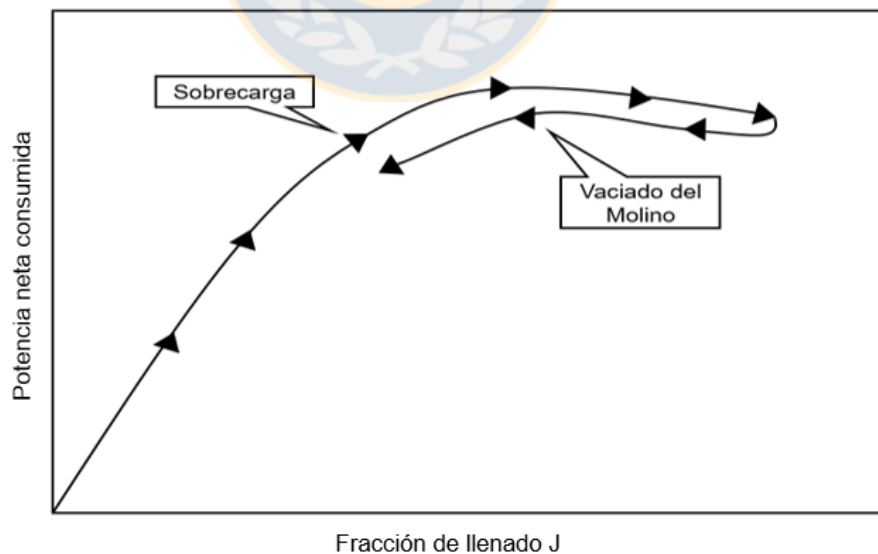


Figura 10: Efecto de la carga fresca (rocas irregulares) y vaciado de una sobrecarga (rocas redondeadas) sobre la demanda de potencia del molino [3]

La Figura 10 muestra la variación de la potencia consumida cuando un molino SAG es alimentado con mineral fresco hasta que se le sobrecarga y luego se evacúa. Es importante destacar que el máximo de la potencia cuando se está cargando el molino con alimentación fresca es mayor que el máximo cuando se encuentra descargando. Este efecto es el resultado del estado (o forma) del mineral. Cuando se agrega alimentación fresca las rocas tienen formas irregulares y presentan un ángulo de reposo mayor. Durante la fase de vaciado los pebbles formados a partir de la alimentación, se habrán redondeado y rodarán más fácilmente sobre la carga, presentando un ángulo de reposo menor y en consecuencia, demandará menos energía.

3.3.2 Distribución Granulométrica en la alimentación

La distribución granulométrica en la alimentación es otro factor que afecta el volumen de la carga en un molino SAG. Operacionalmente se ha mostrado que para un flujo de alimentación constante, el volumen de la carga en el molino es menor, mientras mayor es la proporción de mineral grueso. Visto de otra forma, para un volumen de carga constante, una mayor capacidad de tratamiento se logra cuando el mineral de alimentación es más grueso. Esto ocurre debido a que las colpas de mayor tamaño en la alimentación constituyen un medio de molienda que determina la capacidad moledora del molino. Si la cantidad de gruesos en la alimentación no es suficiente, la intensidad de la molienda en el molino será reducida y la capacidad del molino decrecerá.

El análisis anterior es válido, en el caso de que la dureza del mineral sea normal y permanezca constante y los cambios de granulometría correspondan a problemas de segregación natural en el mineral. Si la cantidad o segregación de gruesos en la alimentación aumenta demasiado, la capacidad de tratamiento del molino disminuirá, ya que estaría faltando la fracción fina que principalmente se muele, es decir el mineral se comportaría como si fuera más duro (menor aptitud para ser molido en forma semiautógena) y la capacidad de tratamiento del molino bajaría. Esto será también así si el mineral se vuelve completamente fino.

Ante variaciones en la distribución de tamaño del mineral alimentado se tiene que decrementos en el mineral de tamaño intermedio y aumento en la misma proporción del mineral más grueso, provocan que el flujo de mineral de descarga aumente, disminuyendo el mineral contenido en el molino, por ello un aumento en el flujo de mineral en la alimentación traerá consigo un aumento en el rendimiento del molino. En cambio ante un aumento del mineral de tamaño intermedio y disminución en la misma proporción del mineral grueso trae consigo una disminución de los medios moledores por lo que la eficiencia del molino se verá disminuida.

La eliminación del material de tamaño intermedio de la alimentación del molino permite aumentar la capacidad de procesamiento del sistema, por ello para evitar el vaciado del molino, la carga de bolas debe disminuirse permitiendo así alcanzar un tonelaje estándar de mineral fresco, es decir eliminando el tamaño intermedio en la alimentación del molino, la capacidad autógena de molienda supera la tasa de alimentación de mineral y el molino se vacía, efecto que es posible contrarrestar disminuyendo la carga de medios molidores de acero.

3.3.3 Dureza del mineral

Mientras más duro es el mineral, mayor será el tiempo que toma su reducción de tamaño, por esto para un flujo de alimentación constante, el volumen de carga aumentará junto con la dureza del mineral. Si el molino está operando con un tonelaje inferior a su capacidad máxima, al aumentar el volumen de su carga consumirá más potencia y el cambio en la dureza se compensará con un aumento del consumo de energía por tonelada de mineral fresco, sin embargo si el molino está siendo ocupado a su máxima capacidad, un aumento en la dureza del mineral producirá un sobrellenado que sólo podrá ser compensado con una disminución del tonelaje tratado.

Cuando no existen problemas de segregación del mineral de alimentación, las variaciones en la granulometría pueden considerarse indicadoras de la dureza relativa del mineral, correspondiendo al mineral más duro una granulometría con mayor proporción de gruesos.

3.3.4 Carga de Bolas

La carga de bolas influye en la operación de un molino semiautógeno, aumentando la densidad media de la carga y provocando un aumento en la demanda de potencia del molino. Para niveles fijos de carga de bolas, cuando se aumenta el nivel del molino a través de un aumento en la carga de mineral, provoca que la densidad media de la carga disminuya a medida que el nivel de llenado se incrementa. Además la carga de bolas permite reducir el tamaño crítico del mineral evitando la acumulación de mineral que provoca una disminución de la capacidad de molienda.

Cuando el molino opera con bajo nivel de carga de bolas, se produce una disminución de la capacidad de procesamiento de mineral, esto debido a la menor disponibilidad de energía mecánica provista por las bolas de acero para la reducción de tamaño del mineral, lo que se traduce en valores inferiores de las tasas específicas de molienda del molino.

A medida que la carga de bolas es disminuida se provoca una acumulación de material de tamaño intermedio, es decir existe cierta especificidad de los medios molidores de acero para actuar sobre el material de tamaño intermedio, lo que supone una correlación entre el nivel de bolas y las tasas específicas de molienda.

En presencia de un nivel de llenado excesivo con una carga de bolas constante la efectividad de los medios moledores sobre el material de tamaño intermedio disminuye, pues la probabilidad de impacto de una bola sobre estos guijarros disminuye al aumentar el nivel de llenado total y mantener constante la carga de bolas, facilitando esto la acumulación de material intermedio durante una operación de elevado nivel de llenado interno del molino.

El volumen de las bolas se expresa como una fracción del volumen del molino.

3.3.5 Densidad y viscosidad de la pulpa

En el molino ocurre que la densidad en la descarga y en el interior del molino no son las mismas esto es debido a que la retención de agua en el molino es generalmente menor que en la de los sólidos finos, por ello que la densidad de la pulpa en el interior sea mayor que en la descarga.

A través de la densidad de la pulpa en la descarga, es posible controlar el nivel de la pulpa en el molino. Aumentando la densidad, se incrementa la viscosidad y se reducen las descargas, provocando un aumento del volumen de la pulpa y de la potencia, además de una disminución de la capacidad de tratamiento de mineral.

El aumento de la potencia se debe a un leve crecimiento de la masa en el molino y del ángulo de apoyo de la carga. Una pulpa más densa y viscosa favorece un ángulo de apoyo mayor lo que significa una mayor demanda de potencia. Cuando el porcentaje de sólidos en la alimentación se mantiene constante la operación estable del molino es solo factible en un reducido margen de tonelajes de mineral en la alimentación, saliendo de este margen, el molino entra en zonas inestables de vaciado y sobrellenado, restringiendo así la manipulación de la variable operación flujo de mineral fresco alimentado.

Los resultados de pruebas en [1] y [2] correspondientes a variaciones del porcentaje de sólidos en la alimentación, muestran un amplio recorrido del nivel de llenado, demostrando la importancia de la proporción de agua en la alimentación del molino sobre el control de las características de transporte de masa al interior del molino. A menor porcentaje de sólidos en la descarga, es decir, mayor cantidad de agua en la alimentación, el nivel de llenado del molino disminuye.

Para el sistema de descarga se deduce que la reología de la pulpa interna define la eficiencia con la cual los alzadores de pulpa retiran el material fino desde el molino. Básicamente, si la viscosidad de la pulpa aumenta, la probabilidad de que ella retorne al molino, una vez superada la parrilla interna, disminuye, pues la capacidad de escurrir de regreso hacia la cámara de molienda es menor con el aumento de viscosidad, respecto de la rapidez con la cual el alzador eleva y descarga el material.

3.4 Modelos de potencia

La predicción del consumo de potencia de un molino SAG está ligada a modelos que combinan factores operacionales y constantes de sintonización. Dentro de estos modelos se destacan Austin, Morrell y Hogg & Fuerstenau [8]. El primero pretende ser un balance de energía cinética y potencia, el segundo corresponde a un balance de la fricción ejercida entre capas concéntricas mientras que el tercero asume que la carga toma la forma de la Figura 6 y que el torque ejercido por ella es igual al que ejercería un punto ubicado en el centro de masa de la carga.

Por motivos de información el único modelo a realizar corresponde a Austin, a continuación se describe la ecuación del modelo de Austin.

3.4.1 Modelo de Austin

$$P_N = K * D^{2,5} * L * (1 - A * J_{total}) * \left[(1 - \varepsilon_b) * \left(\frac{\rho_{Sólido}}{w_c} \right) * J_{total} + 0,6 * J_{bolas} * \left(\rho_{bolas} - \frac{\rho_{Sólido}}{w_c} \right) \right] * N_c * \left(1 - \frac{0,1}{2^{9-10*N_c}} \right) \quad (1)$$

Donde:

P_N : Potencia neta en el piñón de un molino cilíndrico o “mil shell” (kW)

K: Constante de sintonización (Austin propone 10,6)

D: Diámetro del molino dentro de los liners (m)

L: Largo efectivo (m)

A: Constante de sintonización (Austin propone 1,03)

ε_b : Porosidad de la cama de bolas y rocas (Austin propone 0,3 como valor fijo)

ρ_i : Densidad del componente i (t/m³)

J_i : Fracción de llenado del componente i

w_c : Fracción en peso de sólido de la pulpa

N_c : Fracción de la velocidad crítica de giro.

Además Austin propone un factor aproximado de 14% de potencia extra por las zonas cónicas del molino mientras que Alex Doll solamente un 5% [8].

3.5 Modelación de Sistemas

Para estudiar un sistema el primer paso es elaborar un modelo de este. Dicho modelo puede ser una representación de la teoría, el cual se obtiene a partir de las leyes que rigen los fenómenos que afectan al proceso, o un modelo determinado en forma empírica, el cual se determina a partir de la observación directa de los resultados obtenidos al excitar el sistema con entradas conocidas.

Se entenderá por sistema al proceso, objeto o conjunto de elementos que se encuentran bajo análisis.

Por modelo se entenderá a cualquier relación analítica o matemática entre las variables medidas de un sistema. Donde los coeficientes asociados a dichas relaciones analíticas son los parámetros del proceso.

Un conjunto de modelos es un grupo de modelos que comparten una misma estructura y el concepto de clase de modelos se asocia para denominaciones generales de modelos.

Con el desarrollo del modelo, lo que se obtiene es una representación del sistema y su propósito es ayudar a entender dicho sistema. El modelo obtenido puede ser una réplica exacta del sistema o una representación que rescate las propiedades más dominantes de este.

3.5.1 Clases de Modelos

De acuerdo a las distintas características que puede presentar un modelo, se puede distinguir las siguientes:

- De caja blanca, caja negra y caja gris

Los modelos de caja blanca son los llamados modelos fenomenológicos y se construyen en base a ecuaciones físicas del sistema, los modelos de caja negra son los llamados modelos empíricos los cuales se construyen a partir de datos de entrada y salida, los modelos grises son similares a los de caja negra pero incluyen parcialmente aspectos fenomenológicos.

- Continuos y Discretos

Los modelos continuos y discretos son aquellos cuyas variables medidas son señales de tiempo continuo y tiempo discreto respectivamente.

- Estáticos y Dinámicos

Los modelos estáticos son aquellos en los cuales su comportamiento depende sólo de los valores de sus entradas en el instante actual. En cambio un modelo dinámico es aquel que evoluciona dependiendo de sus entradas pasadas y presentes.

- Invariantes y variantes en el tiempo

Un modelo es invariante en el tiempo si su dinámica permanece inalterada, a diferencia de un modelo variante en el tiempo, cuya dinámica evoluciona en el tiempo.

Un modelo $M: u(k) * \theta \rightarrow y(k)$ es invariante en el tiempo si y solo si $M(u(k - \tau), \theta) = y(k - \tau)$, con τ un real y θ un vector de parámetros.

- Lineales y no lineales

De acuerdo a la propiedad de linealidad se tiene que un modelo $M: u(k) * \theta \rightarrow y(k)$ es lineal si y sólo si $M(\alpha_1 u_1(k) + \alpha_2 u_2(k), \theta) = \alpha_1 * M(u_1(k), \theta) + \alpha_2 * M(u_2(k), \theta)$, con α_i reales y θ un vector de parámetros.

- Lineales y no lineales en parámetros

Un modelo $M: u(k) * \theta \rightarrow y(k)$ es lineal en parámetros si y sólo si $M(u(k), \theta_1, \theta_2) = \theta_1 f_1(u(k)) + \theta_2 f_2(u(k))$, con f_1 y f_2 funciones no lineales de $u(k)$.

Nótese que la linealidad en parámetros es una característica que involucra la estructura de un modelo y no debe de confundirse con la propiedad de linealidad que se relaciona con la respuesta del modelo a distintas entradas.



3.5.2 Redes Neuronales Artificiales

El cerebro humano es un sofisticado procesador de información. Se estima que posee aproximadamente 10^{11} neuronas y entre 10^{14} - 10^{15} sinapsis (conexiones entre neuronas) [19]. Las redes neuronales artificiales, RNA's, emulan la arquitectura y capacidades de sistemas neuronales biológicos, para así modelar complejos sistemas.

La Figura 11 muestra el modelo matemático de una neurona artificial. Las entradas x_1, x_2, \dots, x_m representan las m-ésimas dendritas de una neurona, mientras que la salida y_i representa el axón de la misma. Para emular la reacción química de excitación de una neurona, se utiliza una función de activación cuyo argumento corresponde a la suma ponderada de las entradas por medio de una función de integración. En la neurona de la figura 11 la función de activación se denota por $a(\cdot)$ y la función de integración por $f(\cdot)$. La ponderación de las distintas entradas se realiza por medio de los pesos, denotados por $w_{i1}, w_{i2}, \dots, w_{im}$, los cuales representan las sinapsis entre distintas neuronas.

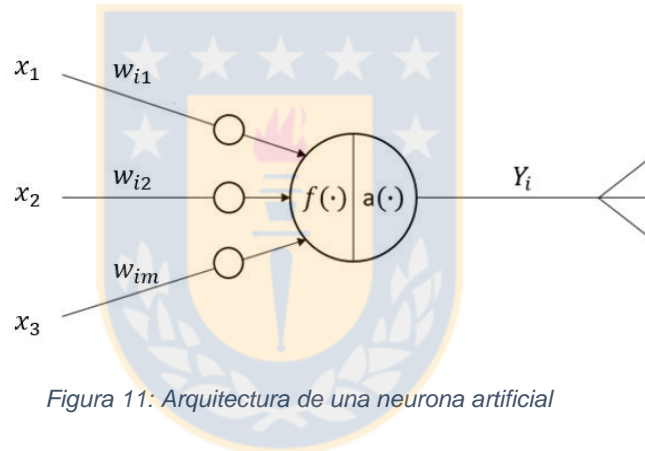


Figura 11: Arquitectura de una neurona artificial

La función de integración es usualmente una función lineal de las entradas, tal como se muestra en la ecuación (2):

$$f_i = \sum_{j=1}^m w_{ij} \cdot x_j - \theta_i \quad (2)$$

Donde,

w_{ij} : j-ésimo peso de ponderación de la i-ésima neurona.

x_j : j-ésima entrada de la i-ésima neurona.

θ_i : Valor umbral de la i-ésima neurona.

Por otro lado, una típica función de activación corresponde a la función sigmoide, cuya ecuación se muestra en la ecuación 3:

$$a(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

Donde,

x : Resultado de la función de integración.

Diferentes funciones de integración y de activación se pueden encontrar en [6]. Cabe mencionar que la función de activación permite que una RNA pueda representar relaciones no lineales [19].

Entre las diversas topologías de RNA's, se encuentran las *single-layer feedforward*, *multilayer feedforward* y *recurrent*. Estas topologías se diferencian en la forma de conexión entre las distintas neuronas [6].

En este trabajo se estudia con mayor detalle la topología *multilayer feedforward* o *perceptrón multicapa* que será utilizada para estimar la potencia. Las RNA's de este tipo, a diferencia de las demás, se componen de varias capas de neuronas interconectadas.

Por otro lado, las neuronas de esta topología solo admiten como entradas las salidas de las neuronas de la capa anterior, lo que las diferencia de las *recurrent networks* que presentan lazos de retroalimentación entre neuronas.

En la Figura 12 se muestra un ejemplo de una red multilayer feedforward:

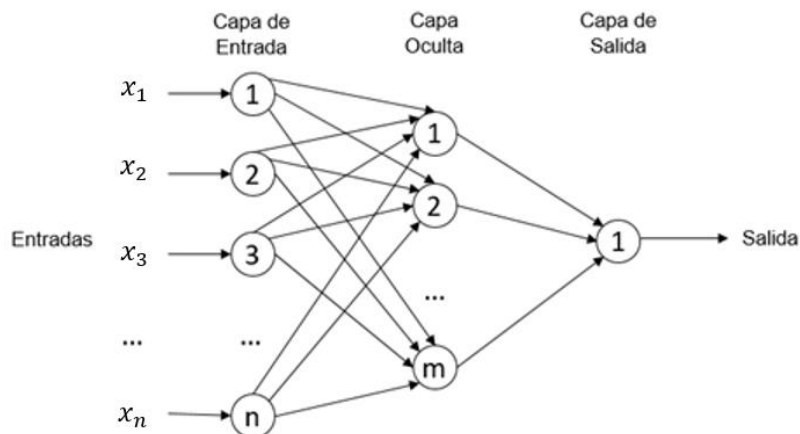


Figura 12: RNA multilayer feedforward

Se distinguen tres tipos de capa:

- Capa de entrada: Está formada por n unidades (siendo n el número de entradas externas) que distribuyen las señales de entrada a la capa siguiente.
- Capas ocultas: Están formadas por neuronas que no tienen contacto físico con el exterior. No existe una restricción para el número de capas ocultas.
- Capa de salida: Está formada por m neuronas (siendo m el número de salidas externas) cuyas salidas constituyen el vector de salidas externas de la red.

Las neuronas en la capa de entrada no llevan a cabo funciones de procesamiento de señal, sino que solo proporcionan un medio por el cual los datos son introducidos a la RNA. Luego, estas señales son traspasadas por medio de las conexiones entre las capas ocultas hasta llegar a la capa de salida.

Por otra parte, la capacidad de aprendizaje de una RNA se encuentra relacionada con la actualización de los valores de los pesos que interconectan las distintas neuronas.

En general, los algoritmos de aprendizaje se clasifican en algoritmos de aprendizaje supervisado, aprendizaje reforzado y aprendizaje no-supervisado.

En el aprendizaje supervisado, que es el utilizado en este trabajo, una red neuronal es alimentada con una secuencia de ejemplos de pares de entrada-salida. El error generado entre la salida de la red neuronal y la salida deseada se utiliza para corregir los valores de los pesos de forma tal que la salida de la red se ajuste a la salida de referencia.

Los tipos de aprendizaje reforzado y no-supervisado se presentan en detalle en [6]

3.5.3 Bosques Aleatorios

3.5.3.1 Árboles de Clasificación

Dentro de los métodos de aprendizaje supervisado, se encuentran los árboles de decisión que corresponde a un modelo de predicción donde dado un conjunto de datos se fabrican diagramas de construcciones lógicas, muy similares a los sistemas de predicción basados en reglas, que sirven para representar y categorizar una serie de condiciones que ocurren de forma sucesiva, para la resolución de un problema [16]. En la Figura 13 queda representada el esquema típico de un árbol de clasificación.

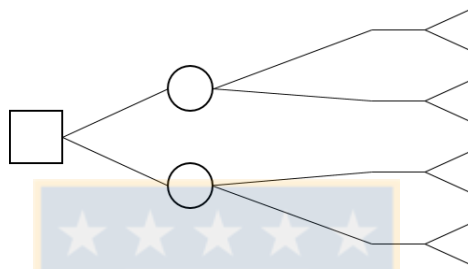


Figura 13: Representación de un árbol de decisión

Se distinguen tres tipos de nodos:

- Nodo de decisión: representado por un cuadrado, muestra la decisión que se tomará.
- Nodo de probabilidad: representado por un círculo, muestra las probabilidades de ciertos resultados.
- Nodo terminal: representado por un triángulo, muestra el resultado definitivo de una ruta de decisión.

3.5.3.2 Random Forest

Random Forest es una combinación de árboles predictivos, el cual trabaja con una colección de árboles incorrelacionados y los promedia [11], en el cual se tiene que cada árbol depende de los valores de un vector aleatorio de la muestra de manera independiente y con la misma distribución de todos los árboles en el bosque. La generalización de error para los bosques converge a un límite en cuanto el número de árboles en el bosque sea grande. El error de generalización de un bosque de árboles de clasificación depende de la fuerza de los árboles individuales en el bosque y la correlación entre ellos. El uso de una selección aleatoria de características para dividir cada nodo produce tasas de error que se comparan favorablemente al algoritmo AdaBoost [10], pero son más robustos con respecto al ruido (clasificador fuerte). Estimaciones internas supervisan el error, la fuerza y la correlación. Además, estos se utilizan para mostrar la respuesta al aumento del número de características utilizadas en la división. Los cálculos internos se utilizan, asimismo, para medir la variable de importancia. Las ideas son también aplicables a la regresión.

El elemento común en todos estos procedimientos es que para el k-ésimo árbol se genera un vector aleatorio θ_k , independiente de los últimos vectores aleatorios $\theta_1, \dots, \theta_{k-1}$ pero con la misma distribución; y un árbol se desarrolla usando el conjunto de entrenamiento y de θ_k , lo que resulta en un clasificador donde $h(x, \theta_k)$ es un vector de entrada.

Como se ha señalado líneas arriba, el método Random Forest se basa en un conjunto de árboles de decisión, es decir, una muestra entra al árbol y es sometida a una serie de test binarios en cada nodo, llamados split, hasta llegar a una hoja en la que se encuentra la respuesta. Esta técnica puede ser utilizada para dividir un problema complejo en un conjunto de problemas simples.

En la etapa de entrenamiento, el algoritmo intenta optimizar los parámetros de las funciones de split a partir de las muestras de entrenamiento.

$$\theta_k^* = \operatorname{argmax}_{\theta_j \in \tau_j} I_j \quad (4)$$

Para ello se utiliza la siguiente función de ganancia de información:

$$I_j = H(j) - \sum_{i \in \{1,2\}} \frac{|S_j^i|}{|S_j|} H(S_j^i) \quad (5)$$

Donde S representa el conjunto de muestras que hay en el nodo por dividir, y S^i son los dos conjuntos que se crean de la escisión. La función mide la entropía del conjunto, y depende del tipo de problema que abordamos [4].

En la Figura 14 queda de manera esquematizada la forma de Random Forest con t árboles de decisión.

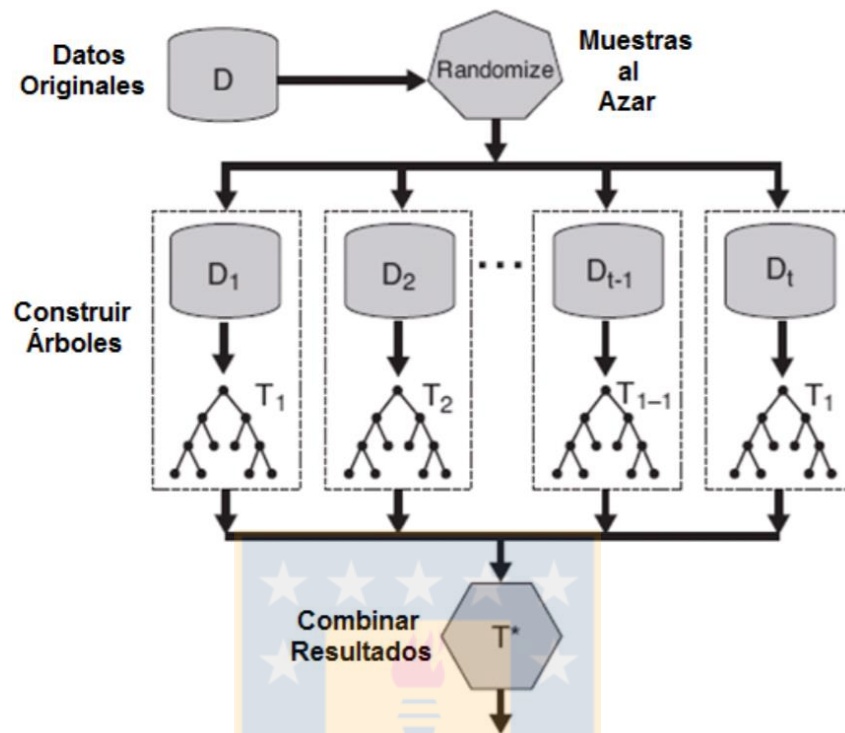


Figura 14: Esquematización de Random Forest

Dentro del modelo de bosques aleatorios se puede determinar la importancia de las variables en la construcción del modelo, esto ayuda a eliminar posibles variables que no generen un aporte, sino, más bien generen ruido al proceso de entrenamiento de los árboles.

Al momento de entrenar cada árbol que se agregará al bosque se deja un tercio de la data de la muestra (data out-of-bag, OOB), con estos datos no utilizados para el entrenamiento se puede estimar un error insesgado de clasificación y también se pueden utilizar para hacer la estimación de importancia de las variables. El funcionamiento de esta lógica corresponde a elegir el error de clasificación out-of-bag, luego se toma una variable al azar y se permutan sus valores dentro de los datos de entrenamiento, ocasionando que dicha variable escogida descorrelacione lo aprendido por el modelo. Luego se vuelve a calcular el error OOB, para luego compararlo con el error calculado inicialmente. En consecuencia, si el error cambia, se afirma que dicha variable es importante, esta importancia queda representada con el nombre de **“Mean Decrease Accuracy”**. Este proceso se repite con todas las variables y luego estas se ordenan de acuerdo a los cambios que produjeron cada una en los errores OOB.

3.6 Validación de los modelos.

Uno de los métodos para la validación de modelos son las funciones de riesgos, como son el error cuadrático medio (MSE), la raíz del cuadrado medio de error (RMSE), el porcentaje de la raíz del cuadrado medio del error (RMSPE), el error absoluto medio (MAE) y el porcentaje de error absoluto medio (MAPE). Todos estos indicadores trabajan sobre los errores entre los valores esperados y los valores observados.

El error cuadrático medio (MSE, por sus siglas en inglés “mean square error”) consiste en las diferencias entre los valores reales Y_1, Y_2, \dots, Y_n y las predicciones del modelo X_1, X_2, \dots, X_n respectivamente [17]. La ecuación 6 muestra el cálculo de esta medida:

$$MSE = \frac{1}{n} \sum_{i=1}^n (X_i - Y_i)^2 \quad (6)$$

Cuando cada par de datos (X_i, Y_i) es mutuamente independiente y el modelo es independiente, el estimado del MSE es una medida confiable de exactitud del modelo. Si los parámetros del modelo fueron ajustados a los datos observados, el MSE subestimaré el verdadero valor de la medida porque el modelo tenderá a reproducir los datos reales [17].

Con frecuencia se utiliza la raíz del cuadrado medio del error (RMSE), mostrada en la ecuación 7 la cual provee una medida del error en las mismas unidades que la variable bajo consideración.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - Y_i)^2} \quad (7)$$

El error absoluto medio (MAE, por sus siglas en inglés “mean absolute error”), mide el error linealmente, como lo muestra la ecuación 8:

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_i - Y_i| \quad (8)$$

Al igual que para el error cuadrático medio, una medida ponderada del error pudiera ser útil, por lo cual se propone el porcentaje del error absoluto medio (MAPE), como se define en la ecuación 9 que provee una medida adimensional robusta [13].

$$MAPE = \frac{1}{n} \sum_{i=1}^n |X_i - Y_i| \times \frac{100}{\bar{Y}} \quad (9)$$

3.6.1 Validación Cruzada (Cross Validation)

La validación cruzada o cross-validation es una técnica utilizada para evaluar los resultados de un análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y prueba. Consiste en repetir y calcular la media aritmética obtenida de las medidas de evaluación sobre diferentes particiones. Se utiliza en entornos donde el objetivo principal es la predicción y se quiere estimar la precisión de un modelo que se llevará a cabo a la práctica [7]. Es una técnica muy utilizada en proyectos de inteligencia artificial para validar modelos generados

3.6.1.1 Validación Cruzada de K iteraciones

En la validación cruzada de K iteraciones, los datos de muestra se dividen en K subconjuntos. Uno de los subconjuntos se utiliza como datos de prueba y el resto ($K - 1$) como datos de entrenamiento. El proceso de validación cruzada es repetido durante k iteraciones, con cada uno de los posibles subconjuntos de datos de prueba. Finalmente se realiza la media aritmética de los resultados de cada iteración para obtener un único resultado. Este método es muy preciso puesto que evaluamos a partir de K combinaciones de datos de entrenamiento y de prueba.

En cada una de las k iteraciones de este tipo de validación se realiza un cálculo de error. El resultado final lo obtenemos a partir de realizar la media aritmética de los K valores de errores obtenidos, según la ecuación (10):

$$E = \frac{1}{K} \sum_{i=1}^K E_i \quad (10)$$

Donde,

K : Número de iteraciones

E_i : Error obtenido en la iteración i -ésima.

3.7 Eliminación de Outliers o Valores Atípicos.

No existe una definición matemática rígida de lo que constituye un valor atípico; determinar si una observación es o no atípica es, en última instancia, un ejercicio subjetivo. Hay varios métodos de detección de valores atípicos [15].

3.7.1 Test de Tukey

El método más impartido académicamente por su sencillez y resultados, que toma como referencia la diferencia entre el primer cuartil y el tercer cuartil, o rango intercuartílico. En un diagrama de caja se considera un valor atípico el que se encuentra 1,5 veces esa distancia de uno de esos cuartiles (atípico leve) o a 3 veces esa distancia (atípico extremo) [15]. En la Figura 15 se muestra el esquema de un gráfico de caja y los valores atípicos.

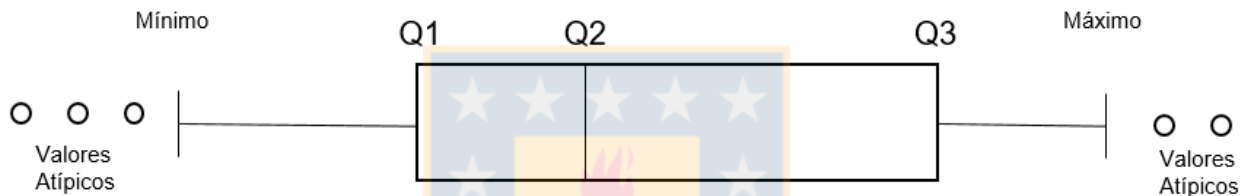


Figura 15: Representación de un gráfico de caja

Dónde,

Q1: Primer cuartil, 25% de los datos es menor o igual que este valor.

Q2: Segundo cuartil, la mediana, 50% de los datos es menor o igual a este valor.

Q3: Tercer cuartil, 75% de los datos es menor o igual que este valor.

4 Desarrollo Experimental

4.1 Variables involucradas en el modelo

Las variables de entrada a considerar en el modelamiento de la potencia del molino SAG, corresponden a variables mineralógicas como lo son la dureza y la granulometría y variables manipulables por el operador que corresponden al porcentaje de sólidos, velocidad de rotación del molino, nivel de bolas y tratamiento y la salida es la potencia del molino SAG, que responde de alguna u otra forma a los cambios en las variables de entrada, las variables quedan representadas en la figura 16:

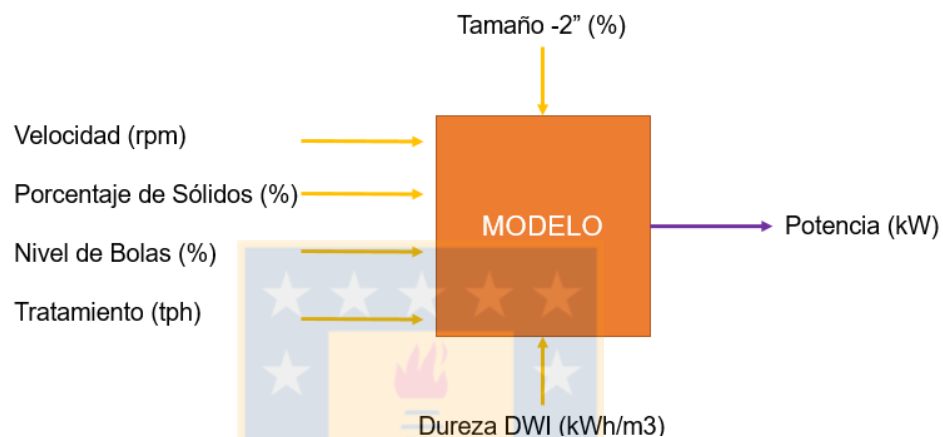


Figura 16: Esquema de variables de entrada y salida en modelo MISO

Contamos con un conjunto de datos de 7 variables, de las cuales 6 de ellas corresponden a las variables independientes de entrada, todas del tipo numérica y una variable dependiente de salida que corresponde a la "Potencia", también del tipo numérica.

4.2 Filtro de data cruda

Bajo condiciones normales de operación del molino SAG 1011 este presenta las siguientes condiciones características:

- Potencia entre 20.000 y 22.100 kW.
- Tratamiento entre 3.500 y 5.000 tph.
- Porcentaje de sólidos entre 67 y 75 %.

Al obtener la data cruda y realizar los filtros correspondientes a valores de operación normales, se utilizaron gráficos de caja e histogramas creados con R para visualizar la presencia de outliers, los cuales quedan representados en las Figuras 17 a 30.

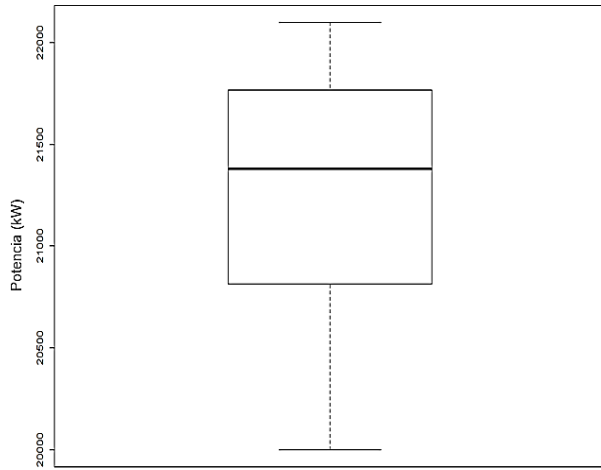


Figura 17: Gráfica de caja para 'Potencia' sin eliminación de outliers

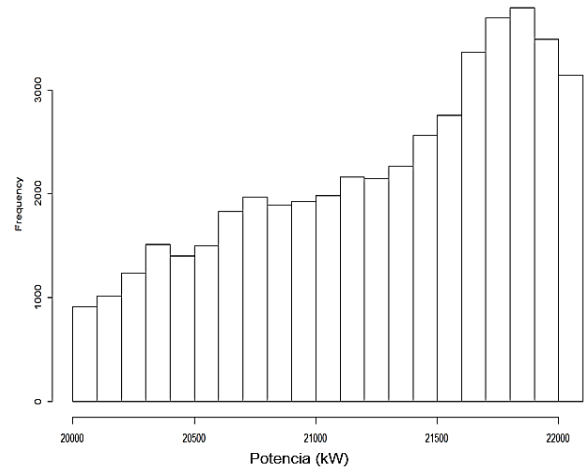


Figura 20: Histograma para 'Potencia' sin eliminación de outliers

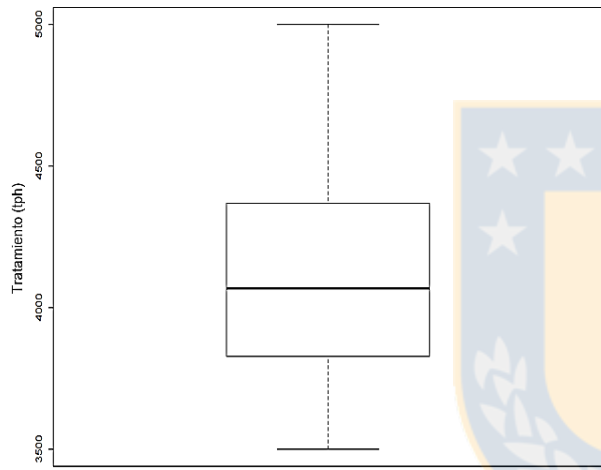


Figura 18: Gráfica de caja para 'Tratamiento' sin eliminación de outliers

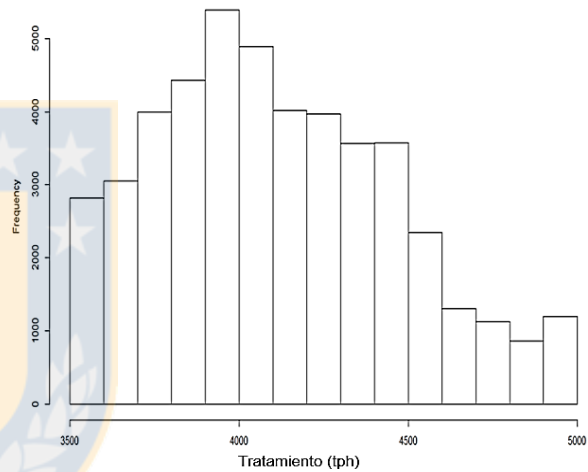


Figura 21: Histograma para 'Tratamiento' sin eliminación de outliers

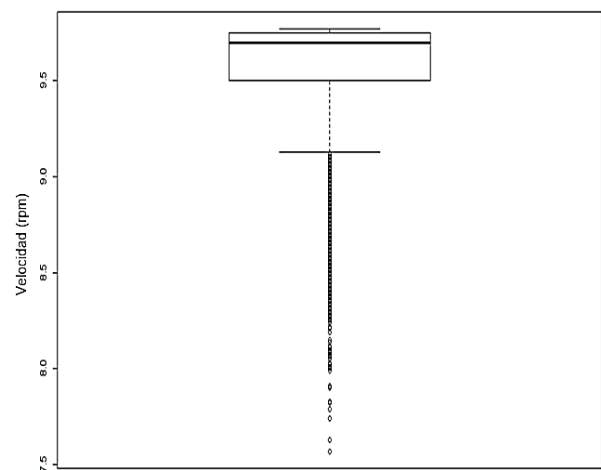


Figura 19: Gráfica de caja para 'Velocidad' sin eliminación de outliers

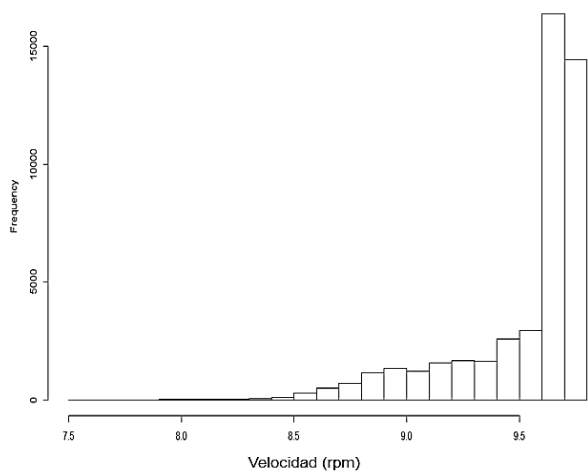


Figura 22: Histograma para 'Velocidad' sin eliminación de outliers

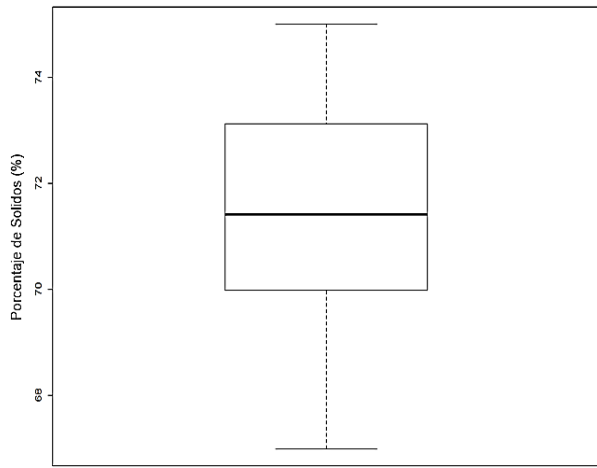


Figura 23: Grafica de caja para 'Porcentaje de Sólidos' sin eliminación de outliers

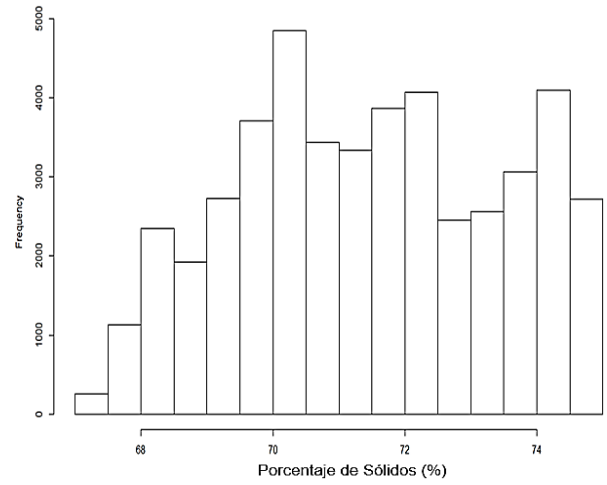


Figura 26: Histograma para 'Porcentaje de Sólidos' sin eliminación de outliers

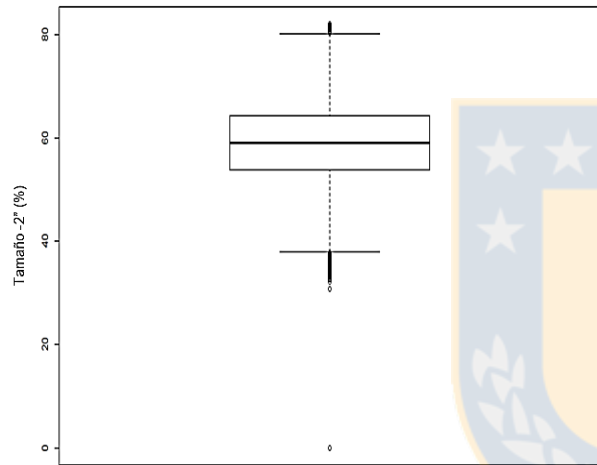


Figura 24: Grafica de caja para 'Tamaño -2\"/>

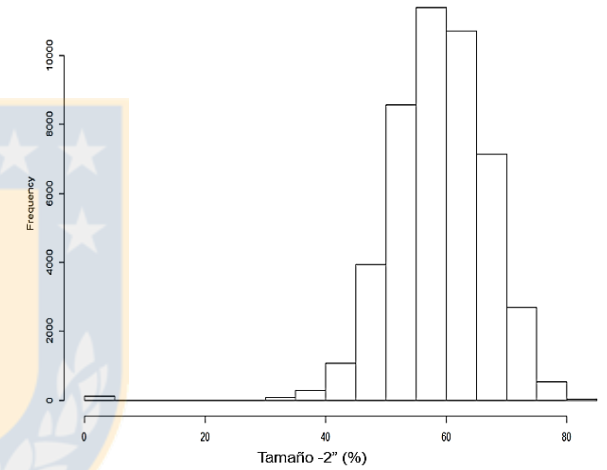


Figura 27: Histograma para 'Tamaño -2\"/>

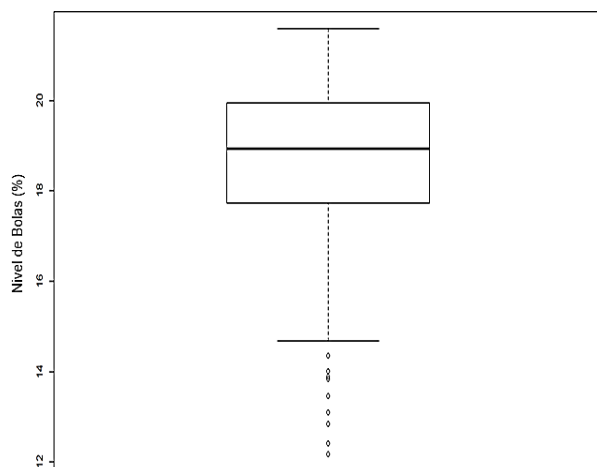


Figura 25: Histograma para 'Porcentaje de Sólidos' sin eliminación de outliers

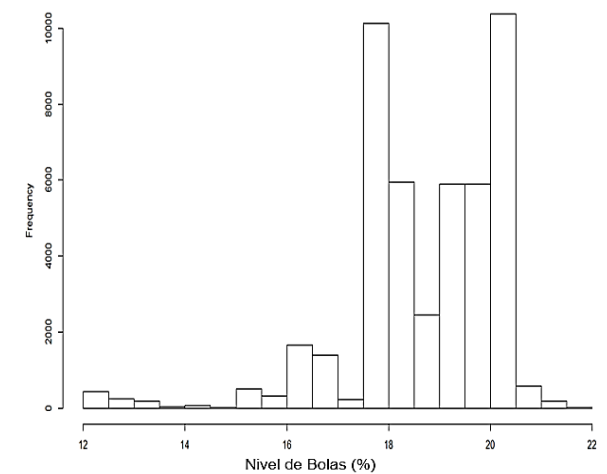


Figura 28: Histograma para 'Nivel de Bolas' sin eliminación de outliers

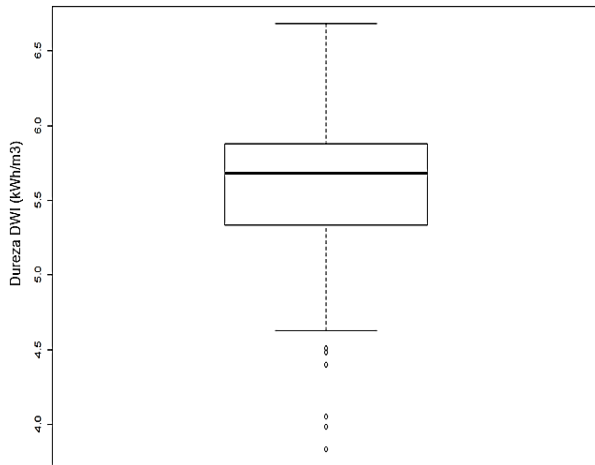


Figura 29: Gráfica de caja para 'Dureza DWI' sin eliminación de outliers

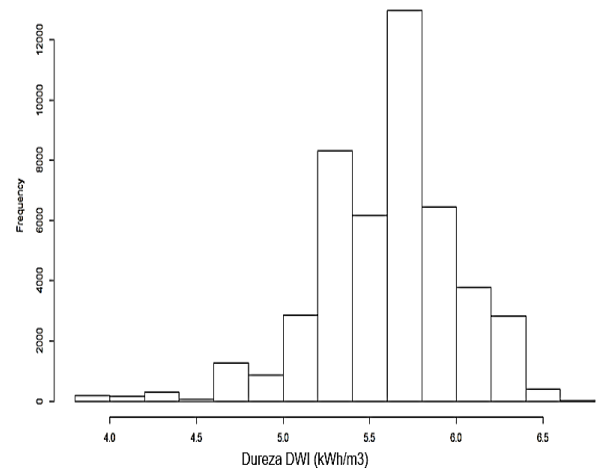


Figura 30: Histograma para 'Dureza DWI' sin eliminación de outliers

En las Figuras 19 y 22 que representan la data en gráfica de caja e histograma respectivamente para la variable de "Velocidad (rpm)" se puede observar la presencia de outliers o valores atípicos desde valores menores a 9,2 (rpm). Lo mismo se observa para las variables "Tamaño -2" (%) donde los valores atípicos se presentan fuera del rango 40-80 (%) en las Figuras 24 y 27, para la variable "Nivel de Bolas (%)" 15-21 en las Figuras 25 y 28 se observan datos atípicos fuera del rango 15-21 (%) y finalmente para la variable "Dureza DWI (kWh/m³)" que se encuentra representada en las Figuras 29 y 30 se observan valores atípicos bajo el valor de 4,9 (kWh/m³).

4.3 Búsqueda de Nivel de Llenado (J) para Modelo de Austin

Actualmente CMDIC no posee un modelo predictivo de nivel de llenado total J ni tampoco posee mediciones regulares de este valor, es debido a esto que fue necesaria la estimación de un nivel de llenado para poder evaluar el desempeño del modelo de Austin.

Con el 70% de la data, la cual está dirigida para el entrenamiento de los modelos, se buscó utilizando la herramienta solver un J constante, que minimizara el error cuadrático medio entre los valores de potencia real y estimados.

Se utilizaron los siguientes parámetros para la ecuación 1:

Tabla 2: Parámetros utilizados para el modelo de Austin

Parámetros Modelo de Austin		
K	-	10,60
D	m	12,20
L	m	6,70
A	-	1,03
ε_b	-	0,30
ρ_{bolas}	t/m ³	7,85

- Se aplicó el modelo de Austin al 70 % de la data para calcular una potencia, dándonos un valor inicial de J total, en este caso se utilizó el 0,35 %. Se calculó el error cuadrático medio entre los valores de potencia reales y los estimados por el modelo de Austin.

	A	B	C	D	E	F	G	H	I	J	K
1	Nivel de Bolas	Velocidad	% Sólidos	Densidad	Potencia Real	Potencia Austin	ECM			K	10,6
2	Jb	Nc	wc	psólido	PR	PA	(PA-PR)^2			D	12,2
3	0,20	0,81	0,75	2,37	20445,92	24308,20	14917195,43			L	6,7
4	0,20	0,80	0,75	2,21	21206,11	23628,50	5867961,93			A	1,03
5	0,19	0,81	0,70	2,37	21607,60	24185,86	6647453,96			EB	0,3
6	0,17	0,78	0,68	2,37	21009,89	22583,58	2476507,28			rho b	7,85
7	0,18	0,80	0,69	2,32	22052,45	23394,51	1801113,64			J	0,35
8	0,18	0,79	0,69	1,81	21962,02	21631,85	109010,76			Error Final	3.625.255
9	0,18	0,79	0,68	1,91	21913,00	22153,72	57944,96				
10	0,20	0,80	0,73	2,07	21678,82	23548,44	3495468,09				
11	0,17	0,80	0,68	2,37	21921,41	23090,29	1366278,15				
12	0,20	0,80	0,73	2,13	21624,76	23711,77	4355583,41				
13	0,20	0,80	0,72	1,93	21276,51	22974,52	2883266,96				
14	0,18	0,80	0,68	2,20	20536,27	23341,04	7866776,08				
15	0,20	0,80	0,72	1,94	22086,93	23186,56	1209193,97				
16	0,18	0,80	0,70	2,32	21938,51	23164,67	1503476,77				
17	0,19	0,80	0,74	2,37	21778,75	23807,91	4117489,92				
18	0,19	0,78	0,74	2,16	21022,67	22457,95	2060034,06				
19	0,19	0,79	0,72	1,81	21605,15	22057,46	204588,84				
20	0,20	0,77	0,72	2,19	20645,49	22778,05	4547776,48				
21	0,19	0,81	0,70	2,22	21693,93	23575,40	3539937,09				
22	0,16	0,80	0,70	2,37	20164,56	22762,69	6750274,91				
23	0,19	0,75	0,71	2,16	20504,20	22035,79	2345782,22				

Figura 31: Extracto de Excel utilizado para el cálculo de nivel de llenado total

- Se utilizó la herramienta solver, seleccionando como celda objetivo el error cuadrático medio con el objetivo de minimizarlo cambiando la variable J.

Parámetros de Solver

Establecer objetivo:

Para: Máx Mín Valor de:

Cambiando las celdas de variables:

Sujeto a las restricciones:

Convertir variables sin restricciones en no negativas

Método de resolución:

Método de resolución
 Seleccione el motor GRG No lineal para problemas de Solver no lineales suavizados. Seleccione el motor LP Simplex para problemas de Solver lineales, y seleccione el motor Evolutionary para problemas de Solver no suavizados.

Ayuda

Figura 32: Solver utilizado para encontrar el valor de J que minimiza el error.

4.4 Búsqueda de la distribución de J en Modelo de Austin

Para encontrar los valores de nivel de llenado total para cada dato utilizado en el modelo, se utilizó la herramienta solver en VBA, que minimizó el error entre la potencia real y la potencia de Austin cambiando el valor de J total.

	B	C	D	E	F	G	H	I
1	Nivel de Bolas	Nc	Sólidos	Densidad	Potencia Real	Potencia Austin	Nivel de llenado	ECM
2	Jb	Nc	wc	psólido	PR	PA	J total	(PA-PR)^2
3	0,20	0,81	0,75	2,37	20445,92	22992,68	0,50	6485952,63
4	0,20	0,80	0,75	2,21	21206,11	22140,74	0,50	873530,42
5	0,19	0,81	0,70	2,37	21607,60	23174,58	0,50	2455444,46
6	0,17	0,78	0,68	2,37	21009,89	22006,39	0,50	993008,68
7	0,18	0,80	0,69	2,32	22052,45	22552,25	0,50	249799,10
8	0,18	0,79	0,69	1,81	21962,02	20102,75	0,50	3456886,78
9	0,18	0,79	0,68	1,91	21913,00	20777,05	0,50	1290384,60
10	0,20	0,80	0,73	2,07	21678,82	21940,64	0,50	68550,05
11	0,17	0,80	0,68	2,37	21921,41	22493,40	0,50	327168,37
12	0,20	0,80	0,73	2,13	21624,76	22186,41	0,50	315445,58
13	0,20	0,80	0,72	1,93	21276,51	21254,67	0,50	476,76
14	0,18	0,80	0,68	2,20	20536,27	22317,01	0,50	3171052,18
15	0,20	0,80	0,72	1,94	22086,93	21449,14	0,50	406777,41
16	0,18	0,80	0,70	2,32	21938,51	22288,02	0,50	122160,34
17	0,19	0,80	0,74	2,37	21778,75	22676,16	0,50	805341,83
18	0,19	0,78	0,74	2,16	21022,67	21116,15	0,50	8737,71
19	0,19	0,79	0,72	1,81	21605,15	20315,27	0,50	1663786,45
20	0,20	0,77	0,72	2,19	20645,49	21487,83	0,50	709532,12
21	0,19	0,81	0,70	2,22	21693,93	22414,01	0,50	518516,43
22	0,16	0,80	0,70	2,37	20164,56	22125,93	0,50	3846987,65
23	0,19	0,75	0,71	2,16	20504,20	20841,12	0,50	113517,92

Figura 33: Extracto de Excel utilizado para encontrar la distribución de nivel de llenado total

Luego grabamos la macro para que resuelva cada punto minimizando el error en la columna "I" y variando J total que se encuentra en la columna "H".

```

Project - VBAProj (X)
  (General)
  Sub MultipleSolver()
  Dim i As Integer
  ActiveWorkbook.ActiveSheet.Activate
  For i = 2 To 26035
  SolverReset
  SolverOk SetCell:="$I$" & i, MaxMinVal:=2, ByChange:="$H$" & i

  SolverSolve True
  Next i
  End Sub
  
```

Figura 34: Macros utilizadas para solución de solver

5 Resultados

5.1 Eliminación de Outliers

Una vez eliminados los valores entregados por los gráficos de caja, se vuelve a graficar, obteniendo diagramas sin presencia de valores atípicos. En las Figuras 35 a 48 se observan las gráficas de caja e histogramas de las 7 variables utilizadas para el modelamiento, donde se observa que no hay presencia de valores atípicos en la muestra.

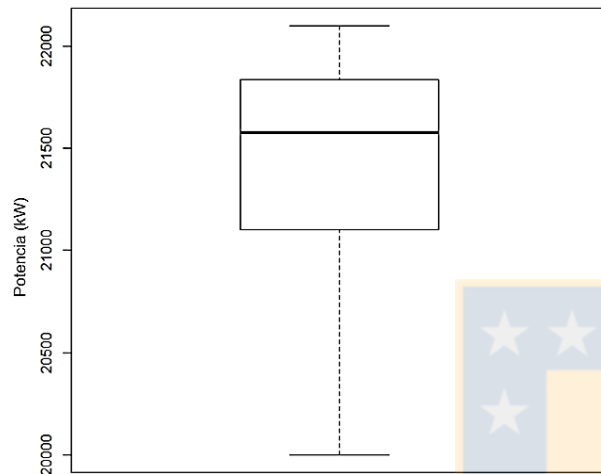


Figura 35: Gráfica de caja para 'Potencia' sin presencia de outliers

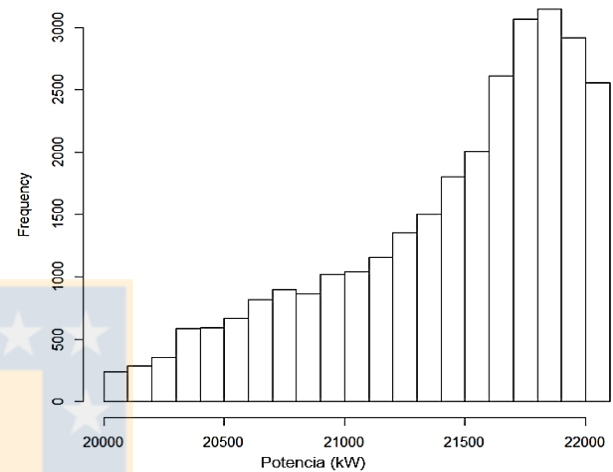


Figura 37: Histograma para 'Potencia' sin presencia de outliers

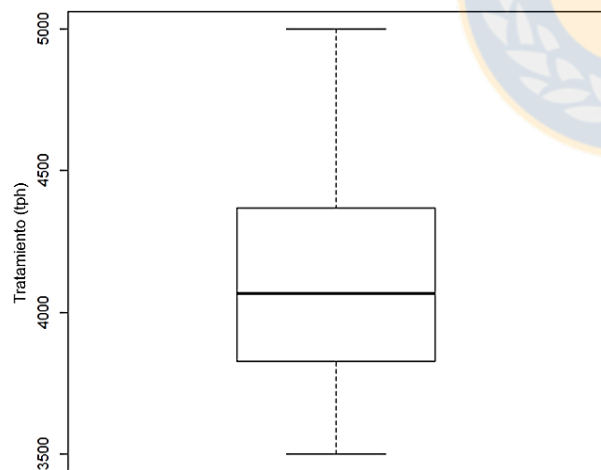


Figura 36: Gráfica de caja para 'Tratamiento' sin presencia de outliers

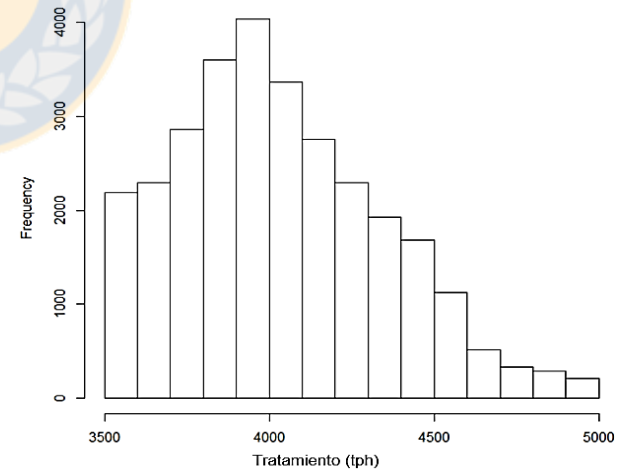


Figura 38: Histograma para 'Tratamiento' sin presencia de outliers

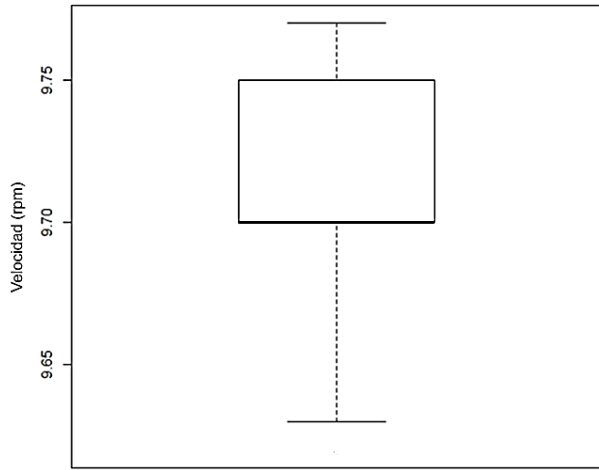


Figura 39: Gráfica de caja para 'Velocidad' sin presencia de outliers

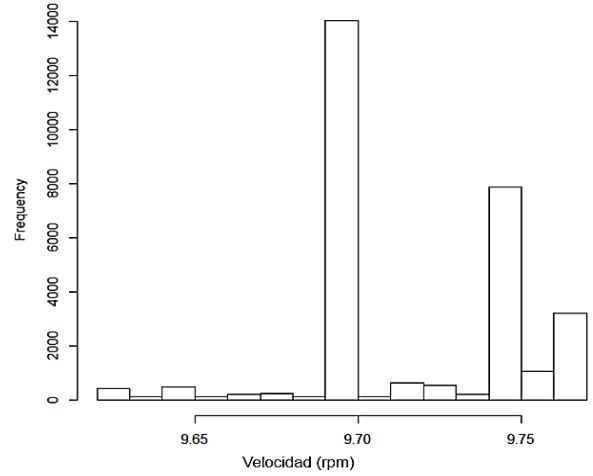


Figura 42: Histograma para 'Velocidad' sin presencia de outliers

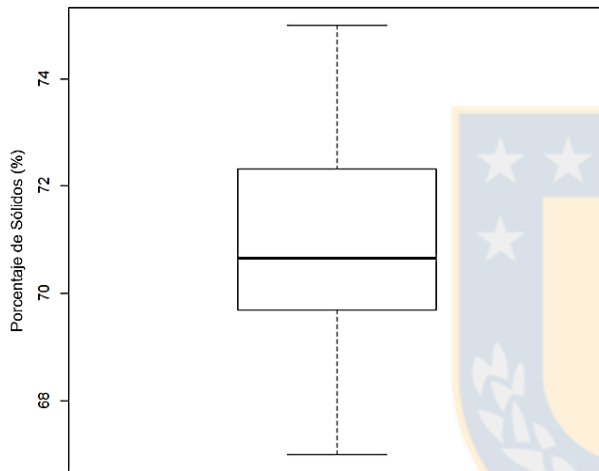


Figura 40: Gráfica de caja para 'Porcentaje de Sólidos' sin presencia de outliers

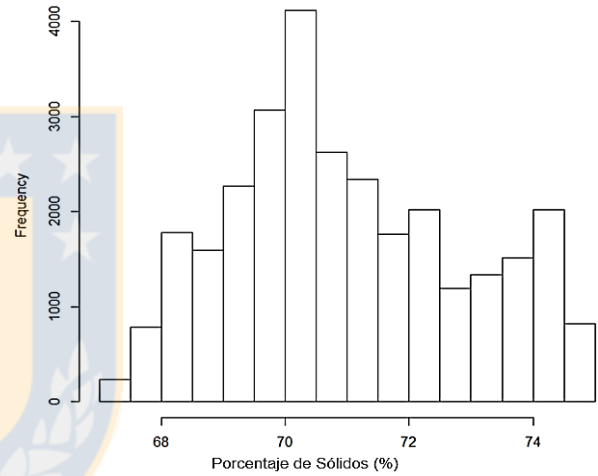


Figura 43: Histograma para 'Porcentaje de Sólidos' sin presencia de outliers

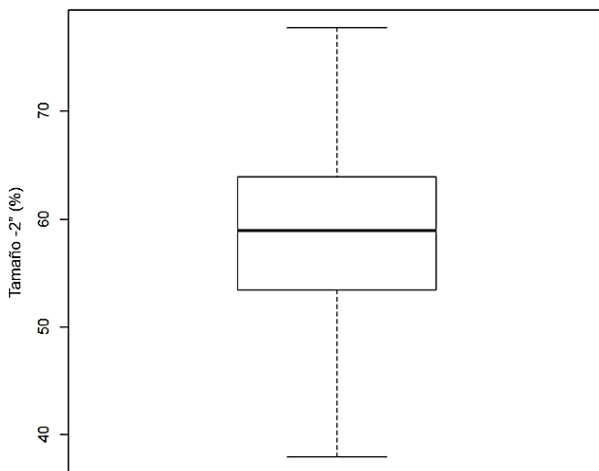


Figura 41: Gráfica de caja para 'Tamaño -2" sin presencia de outliers

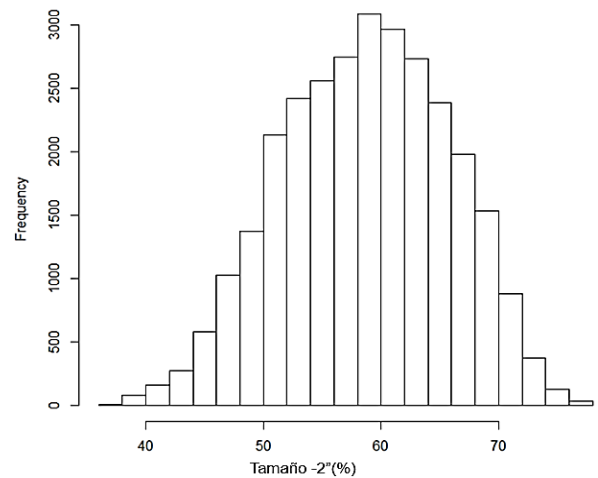


Figura 44: Histograma para 'Tamaño -2" sin presencia de outliers

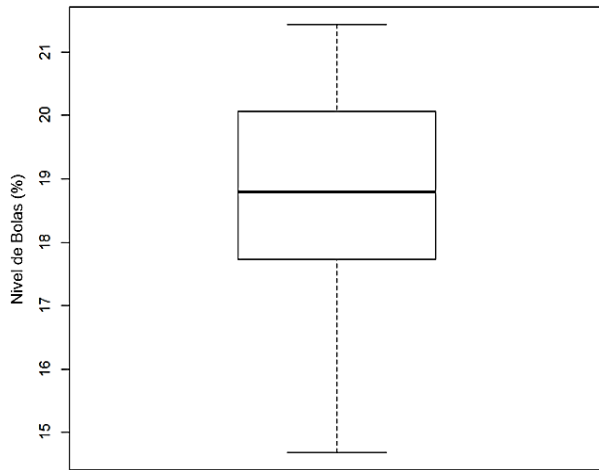


Figura 45: Gráfica de caja para 'Nivel de Bolas' sin presencia de outliers

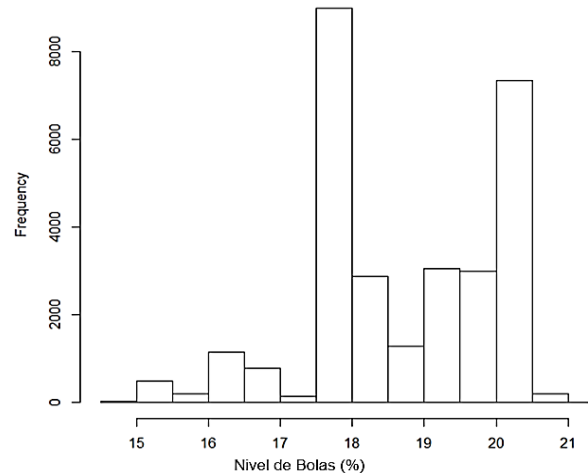


Figura 47: Histograma para 'Nivel de Bolas' sin presencia de outliers

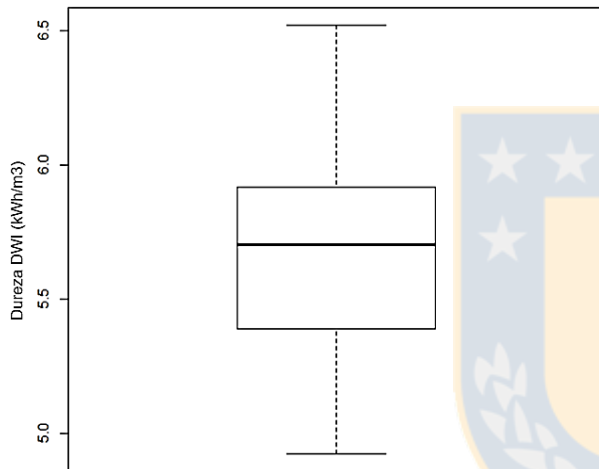


Figura 46: Gráfica de caja para 'Dureza DWI' sin presencia de outliers

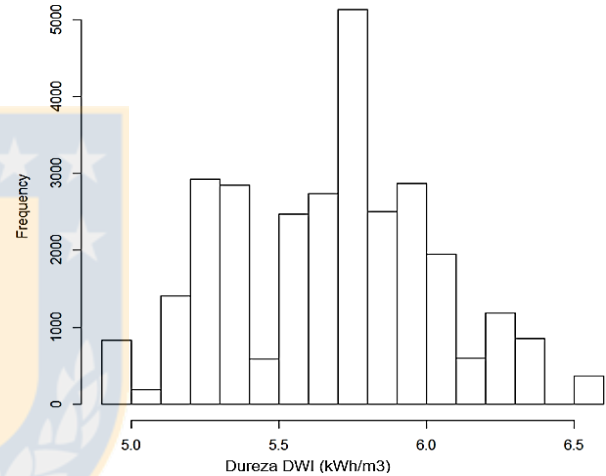


Figura 48: Histograma para 'Dureza DWI' sin presencia de outliers

Algunos valores de interés que demuestran que los datos se encuentran dentro de rangos operacionales normales son mínimo, máximo, media y desviación estándar, las cuales quedan representados en la Tabla 3:

Tabla 3: Resumen de las variables numéricas en el conjunto de datos disponibles

Parámetro	Unidad	Mínimo	Máximo	Promedio	Desviación Estándar
Tratamiento	tph	3.500,00	4.980,87	4.032,27	314,47
Velocidad	rpm	9,62	9,77	9,72	0,03
%Sólidos	%	67,00	75,00	71,00	1,89
% Tamaño -2"	%	37,94	77,79	58,67	7,09
Nivel de Bolas	%	14,68	21,44	18,71	1,29
Dureza DWI	kWh/m ³	4,92	6,52	5,67	0,35
Potencia	kW	20.000,04	22.099,94	21.431,96	511,45

5.2 Generación de modelos con Redes Neuronales

Se generarán modelos de redes neuronales perceptrón multicapa, considerando como variables de entrada el Tratamiento, la Velocidad, el Nivel de Llenado de Bolas, la Granulometría, la Dureza y el Porcentaje de Sólidos, mientras que la variable de salida la potencia (Ver anexo 8.2). Los parámetros a considerar para las distintas pruebas se muestran en la Tabla 4.

Tabla 4: Parámetros a considerar para la generación de los modelos RNA's

Datos para entrenamiento	70%
Datos para validación	30%
Número de neuronas	1 a 10

- Todos los modelos serán creados, entrenados y validados usando la arquitectura Feedforward y el algoritmo de entrenamiento Perceptrón Multicapa con una capa de entrada, una capa oculta y una de salida.
- La importancia de destinar el mayor porcentaje de datos para entrenamiento radica en que la red neuronal puede aprender más al haber más puntos de operación, por lo tanto, será un valor que se mantendrá en todas las pruebas, al igual que los porcentajes de validación.
- En primera instancia se probará con un intervalo de 1 a 10 para el número de neuronas en la capa oculta, obteniéndose 10 modelos en total, dependiendo de los resultados y el análisis que pueda obtenerse de ellos, se probará con aumentar dicho intervalo.
- El criterio para identificar la mejor red neuronal es seleccionar los RSME de validación más bajo, como prueba de que dio buenos resultados con un nuevo set de datos.

Tabla 5: Número de neuronas versus errores RSME de entrenamiento, validación y validación cruzada con $k = 10$ iteraciones

Número de Neuronas	RSME Entrenamiento (kW)	RSME Validación (kW)	RSME Validación Cruzada (kW)
1	457,99	458,19	463,54
2	442,17	445,41	450,07
3	461,17	463,23	474,73
4	475,98	480,10	489,15
5	496,23	498,70	511,82
6	519,04	520,13	539,46
7	540,36	542,02	547,83
8	526,38	527,56	542,74
9	562,93	566,86	575,48
10	590,03	592,81	603,47

De acuerdo a los resultados obtenidos de la Tabla 5, los mejores errores de validación y test se obtienen para las redes neuronales con 2 neuronas en la capa oculta. La Figura 49 muestra la raíz de errores cuadráticos medios de entrenamiento y validación para redes variando el número de neuronas.

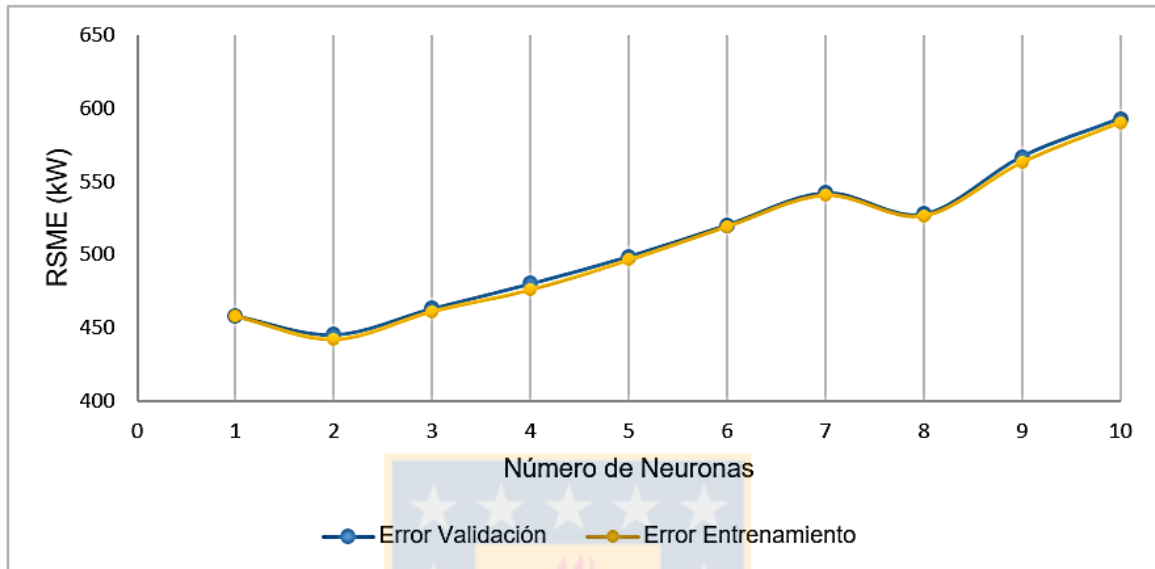


Figura 49: Raíz de error cuadrático medio de entrenamiento y validación cambiando el número de neuronas en la capa oculta

En la Figura 49 se observa que ambos errores tienden al alza al aumentar el número de neuronas en la capa oculta. Esto se debe a que los pequeños ajustes de peso pueden modificar el comportamiento de la red completamente, siendo muy fácil perder el control de la red. Así, a mayor número de neuronas, menos fiabilidad tiene la red.

Luego graficando los valores de validación cruzada con $k = 10$ iteraciones, podemos notar que nuevamente la mejor red neuronal fue la que tiene dos neuronas en la capa oculta. Esto queda representado en la Figura 50.

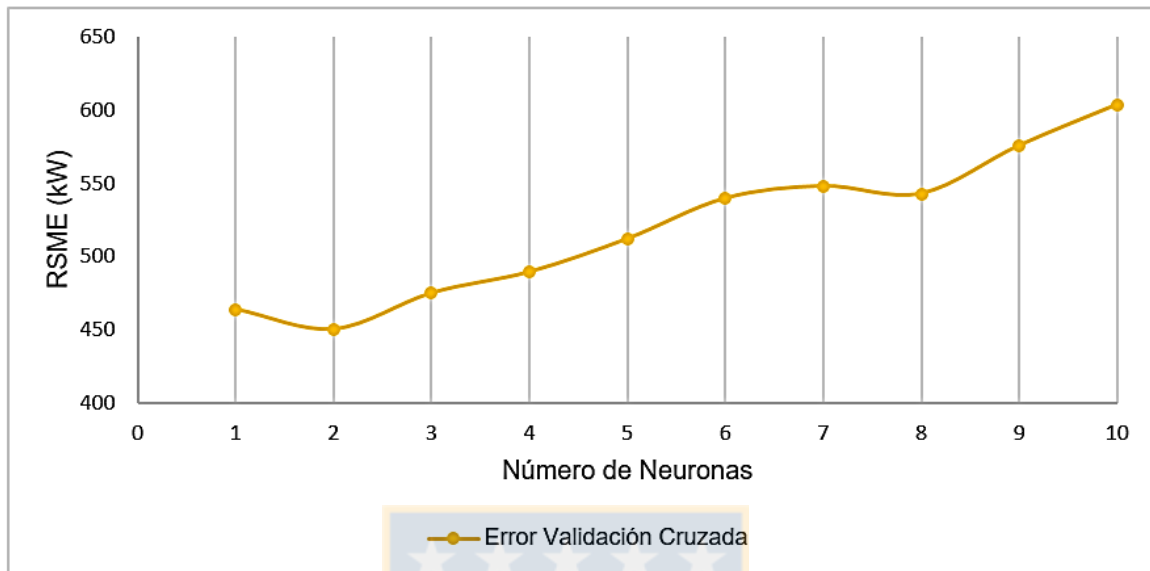


Figura 50: Raíz del error cuadrático medio de validación cruzada cambiando el número de neuronas en la capa oculta

Bajo este análisis podemos corroborar que la mejor red neuronal efectivamente fue la de 2 neuronas en la capa oculta.

Luego, los errores asociados al modelo de redes neuronales con 2 neuronas en la capa oculta se ordenan en la Tabla 6:

Tabla 6: Errores de Red Neuronal con 2 neuronas

Número de Neuronas	RSME Entrenamiento (kW)	RMSE Validación (kW)	MAPE (%)	RMSE Validación Cruzada (kW)
2	442,17	445,41	1,66	450,05

Finalmente la red queda representada como lo muestra la Figura 51:

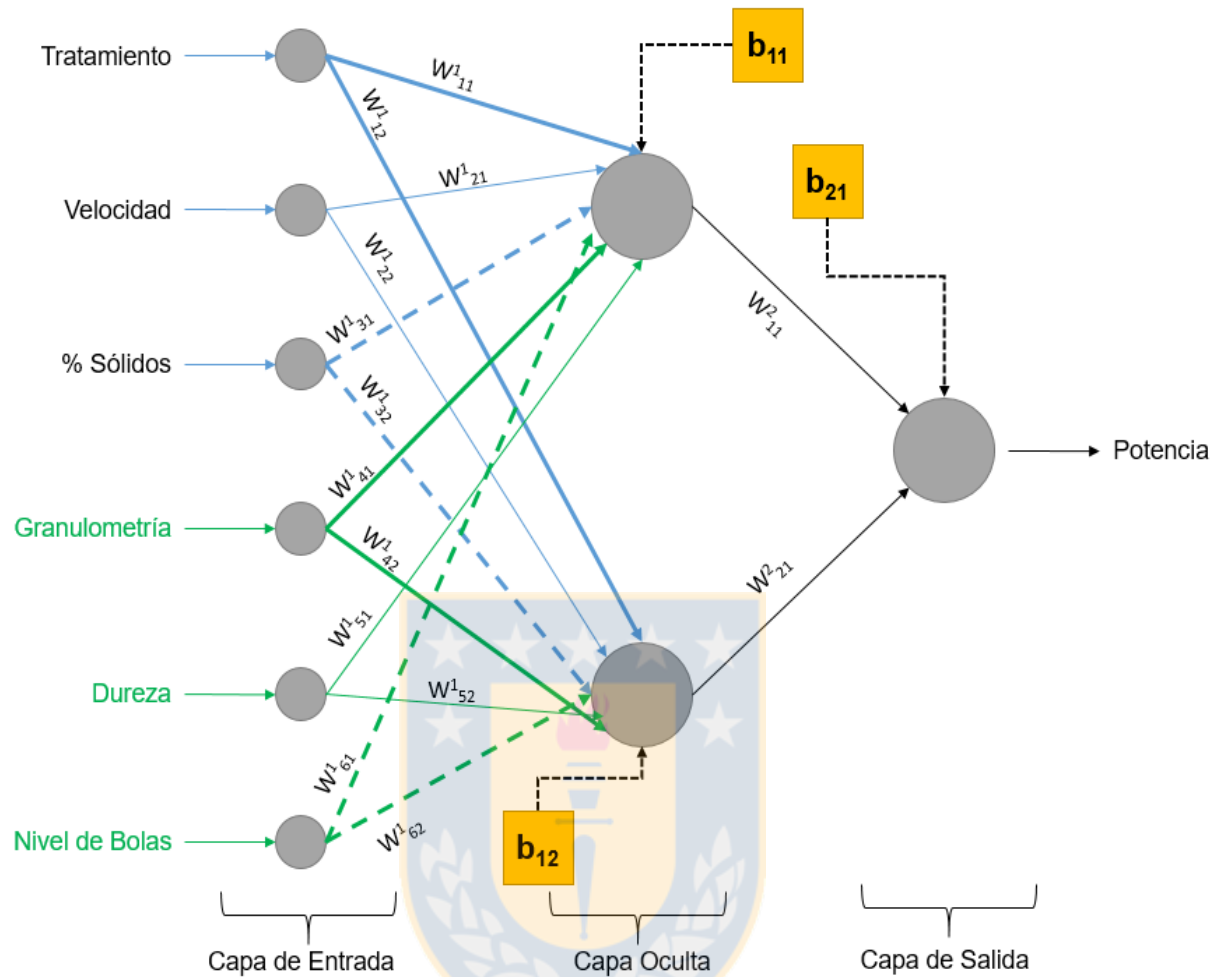


Figura 51: Representación de la red neuronal con 2 neuronas en la capa oculta

En las Tablas 7,8 y 10 se observan los valores de los pesos que interconectan las neuronas de la capa de entrada con la capa oculta, y los pesos que conectan la capa oculta con la capa de salida.

En la Tabla 9 se observan los valores de los pesos sinápticos pertenecientes a la capa Bias.

Tabla 7: Pesos correspondientes a cada conexión entre las entradas y la neurona 1 de la capa oculta

Tratamiento	W^1_{11}	-4,10241
Velocidad	W^1_{21}	-60,46185
% Sólidos	W^1_{31}	4,37958
Granulometría	W^1_{41}	2,83172
Dureza	W^1_{51}	0,33081
Nivel de Bolas	W^1_{61}	12,12389

Tabla 8: Pesos correspondientes a cada conexión entre las entradas y la neurona 2 de la capa oculta

Tratamiento	W^1_{12}	1,38436
Velocidad	W^1_{22}	28,10082
% Sólidos	W^1_{32}	-1,75229
Granulometría	W^1_{42}	-1,97622
Dureza	W^1_{52}	0,26743
Nivel de Bolas	W^1_{62}	-6,01451

Tabla 9: Valores de bias de las neuronas en la capa oculta y capa de salida de la mejor red neuronal

Bias	
b_{11}	20,13842
b_{12}	-8,68711
b_{21}	1,9925

Tabla 10: Pesos que conectan las neuronas de la capa oculta con la capa de salida

Pesos en la capa de salida	
W^2_{11}	-1,27187
W^2_{21}	-1,3353

5.3 Generación de modelo utilizando Random Forest

En los modelos Random Forest la variable a modificar corresponde al número de árboles de decisión (Ver anexo 8.3). Se generarán modelos de bosques aleatorios, considerando como variables de entrada el Tratamiento, la Velocidad, el Nivel de Llenado de Bolas, la Granulometría, la Dureza y el Porcentaje de Sólidos , mientras que la variable de salida la Potencia. Todos los modelos serán entrenados utilizando la librería "RandomForest" del software estadístico R [22].

Los parámetros a considerar para las distintas pruebas se muestran en la Tabla 11:

Tabla 11: Parámetros a considerar para la generación de los modelos random forest

Datos para entrenamiento	70%
Datos para validación	30%
Número de árboles	0 a 500

Se entrena el modelo con un número de árboles iniciales de 500, lo que nos da un resultado del error cuadrático medio como sigue:



Figura 52: Error cuadrático medio (con datos escalados) versus número de árboles

Según la Figura 52 podemos notar que con 150 árboles es suficiente ya que el error comienza a converger.

Luego para un modelo Random Forest con 150 árboles de decisión, los errores asociados son:

Tabla 12: Errores correspondientes a Random Forest con 150 árboles de decisión

Número de Árboles	RMSE Validación (kW)	MAPE (%)	RMSE Validación Cruzada (kW)
150	246,66	0,76	247,19

Gracias a la función 'importance' podemos encontrar las variables que fueron más influyentes para la construcción del modelo según la medida Mean Decrease Accuracy, las cuales se esquematizan en la Figura 53:

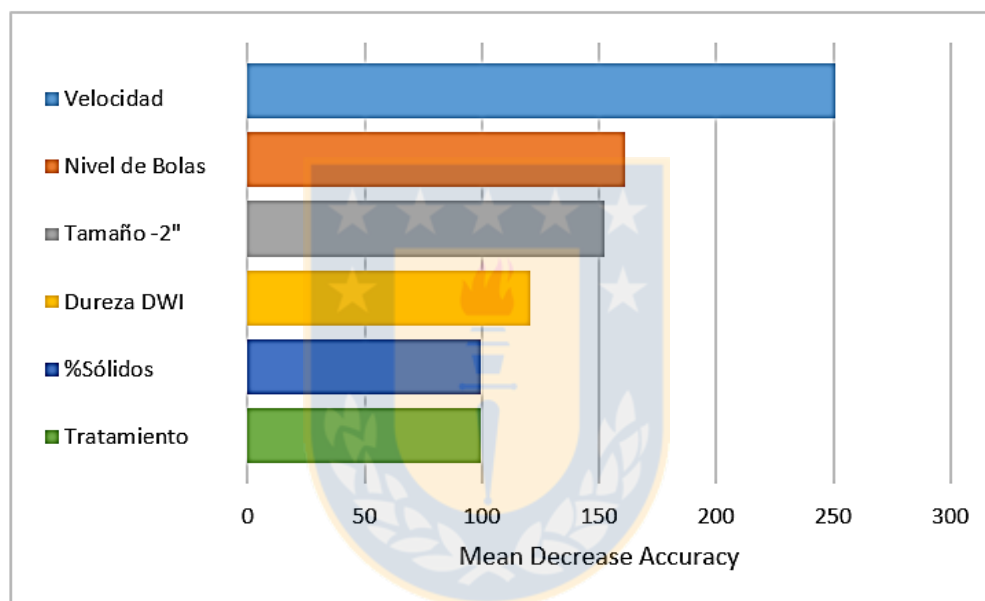


Figura 53: Importancia de las variables en el modelo RF

5.4 Modelo de Austin

Para la aplicación de este modelo el valor de J_{total} encontrado fue de $J_{total} = 52,59\%$.

Luego, los errores asociados utilizando el 30% de la data en el modelo de Austin son:

Tabla 13: Errores asociados al modelo de Austin con J constante

J constante (%)	RMSE (kW)	MAPE (%)
52,59	943,76	3,04

Para analizar qué tan acertado fue el valor constante a elegir, se buscó la distribución de valores de J total para cada dato utilizando solver, estos resultados se muestran en el histograma de la Figura 54:

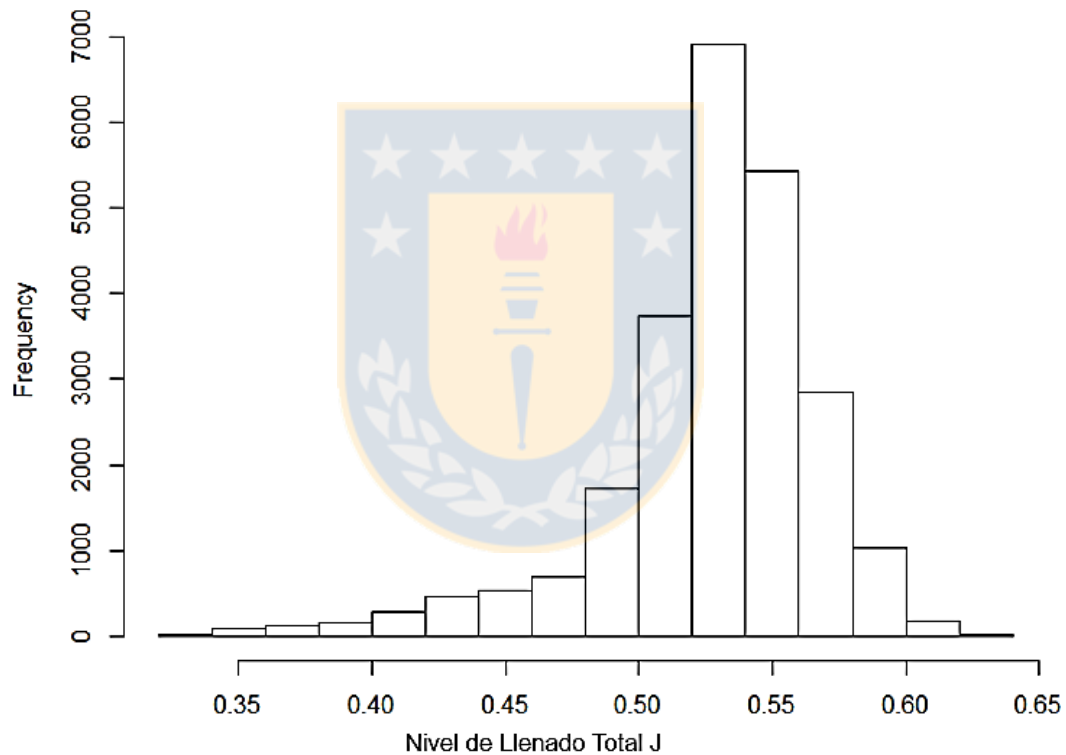


Figura 54: Distribución del nivel de llenado total del molino

Al observar la distribución de J podemos decir que el valor constante seleccionado se encuentra dentro del rango de mayor frecuencia, pero de todas maneras castigamos el desempeño del modelo ya que sus valores ajustados son dispersos.

5.5 Desempeño de Modelos

Considerando el error:

$$error = \hat{y} - y \quad (11)$$

Con,

\hat{y} : Valor de potencia predicho por los distintos modelos

y : Valor real de potencia.

Las Figuras 55, 56 y 57 muestran los histogramas de la distribución del error de los distintos modelos al momento de predecir los valores de potencia utilizando data de validación.

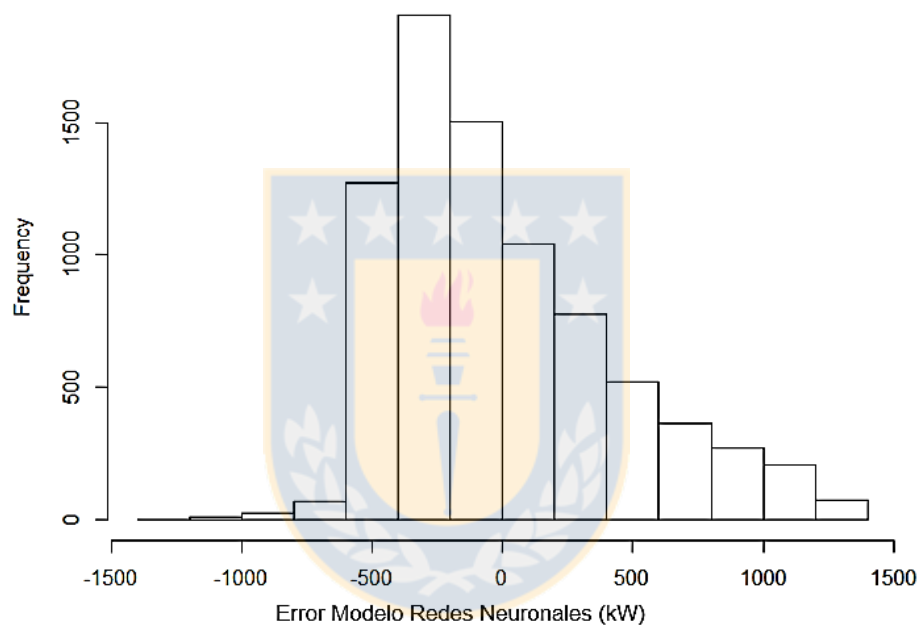


Figura 55: Histograma de error para modelo redes neuronales

La Figura 55 muestra la distribución del error correspondiente al modelo de redes neuronales, esta toma valores desde -1.500 (kW) hasta 1.500 (kW). Existe una alta frecuencia para el valor de 0 (kW), sin embargo, el peak se alcanza para valores cercanos a -400 (kW). Además, se puede comentar que la mayor concentración de datos se encuentra para valores entre 0 (kW) y -500 (kW), esto indica que el modelo tiende a entregar valores de potencia menores a los reales.

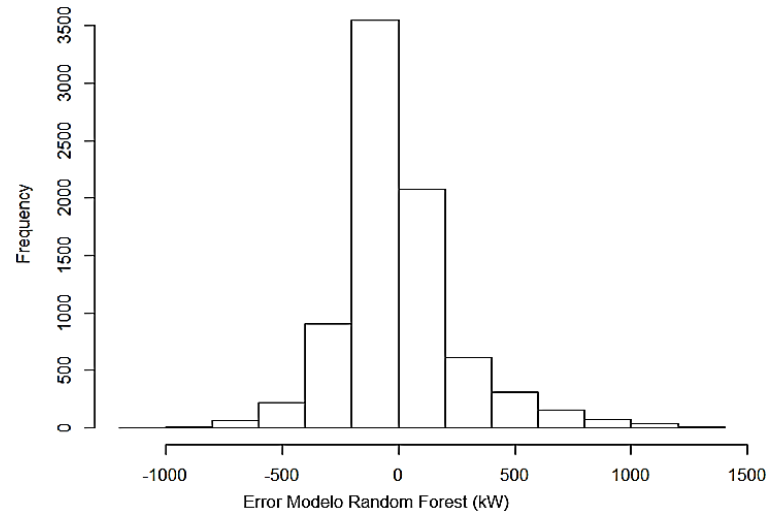


Figura 56: Histograma de error para modelo random forest

La Figura 56 muestra el histograma de la distribución del error para el modelo de random forest, este toma valores desde -500 (kW) a 500 (kW). El peak de la frecuencia se encuentra para el valor de -200 (kW), basado en esto se puede decir, que el modelo predice en su mayoría valores de potencia menores a los reales.

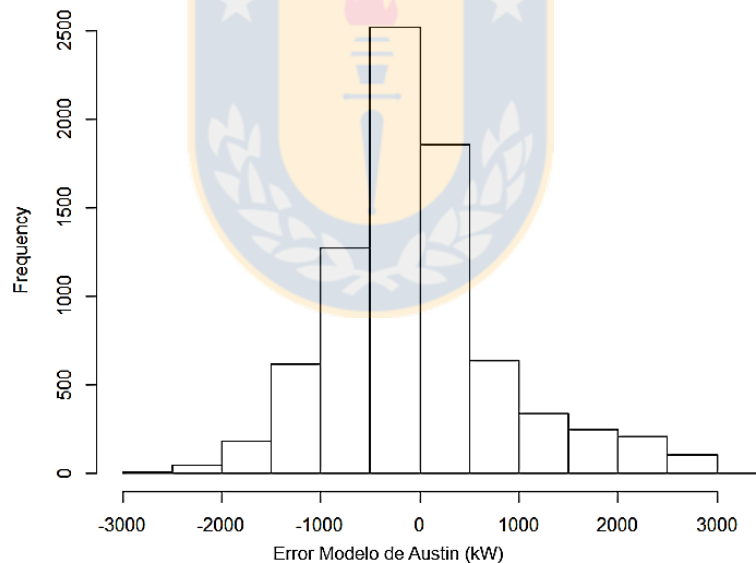


Figura 57: Histograma de error para modelo de Austin

En la Figura 57 se muestra la distribución del error correspondiente al modelo propuesto por Austin [8], este toma valores desde -3.000 (kW) hasta 3.000 (kW). Si bien se observan valores cercanos a 0 (kW), el peak de frecuencia se tiene en el valor aproximado de -500 (kW), generándose nuevamente la mayor concentración de datos en el sector negativo del histograma. El modelo de Austin entregó principalmente valores de potencia menores a los reales.

6 Conclusiones

Se propusieron tres tipos de modelos estáticos con el fin de determinar el mejor que prediga el comportamiento de la potencia del molino SAG 1011 de Minera Collahuasi. De los resultados se concluye que el modelo de bosques aleatorios o random forest con 150 árboles de decisión tiene un mejor desempeño que el modelo de redes neuronales con 2 neuronas en la capa oculta. Mientras los bosques tienen una raíz de error cuadrático medio de 246,66 (kW); el del modelo de redes llega a 445,41 (kW), por otro lado el modelo establecido de Austin tuvo un desempeño de 943,76 (kW).

Con respecto a los modelos random forest, aumentar el número de árboles provocó una disminución del error hasta cierto valor de convergencia, lo cual es esperado. Según la medida Mean Decrease Accuracy la variable más influenciada en el proceso de entrenamiento de los árboles es la velocidad (rpm) de rotación del molino, lo cual no es de extrañar ya que en la vida real, el operador al ver alteraciones en los valores de potencia en el molino su primera acción es variar la velocidad, a valores muy altos de potencia el sistema experto actúa y puede llegar a detener el molino, antes que esto ocurra el operador aumenta las revoluciones por minuto para provocar un mayor flujo de descarga de mineral.

Por otro lado el modelo RNA al aumentar el número de neuronas el error tendió al aumento por lo que se concluye que debido a los pequeños ajustes de peso se modificó el comportamiento de la red, perdiendo el control de ella, así a mayor número de neuronas, menos fiabilidad tiene la red, de todas formas, si en el caso del molino SAG 1011 basta con sólo 2 neuronas en la capa oculta, para otros sistemas puede que se necesiten más. Esto sólo se demuestra con prueba y error. Al observar la figura 49 se observó que el error de entrenamiento es menor que el de validación pero la diferencia entre ellos no es significativamente mayor por lo que no ocurre un sobreajuste u overfitting en el modelo.

Otra gran diferencia al momento de poner en marcha el entrenamiento de los modelos, fue la gran diferencia en la exigencia computacional requerida, cuando el entrenamiento de los bosques tardaba entre 20 a 30 minutos, las redes tardaban un tiempo aproximado de 2 horas el cual fue en aumento al aumentar el número de neuronas en la capa oculta.

No es de extrañar que el modelo de Austin pudiera tener una raíz del error cuadrático medio (RMSE) mayor a los modelos propuestos, ya que depende cuadráticamente del nivel de llenado total del molino, el cual mantuvimos constante. Sin embargo, lo que llama la atención es que el error no sea considerablemente mayor, ya que según lo observado en la figura 54, los valores puntuales de nivel de llenado total J se encuentran dispersos entre el 30% y 60% pero sí mantienen su peak de frecuencia entre el 52% y 54%. Es debido a esto que podemos decir que el modelo de Austin fue castigado al mantener el valor de nivel de llenado constante y podría tener un desempeño mucho mejor si se utilizaran valores de J adecuados. Con respecto a los modelos propuestos podemos destacar que el error no es drásticamente diferente entre ellos y de todas formas, que el error sea más o menos grande no es tan relevante si lo que se quiere es conocer *a priori* los valores que tendrá la potencia más que su valor exacto (partiendo de la base de que ningún modelo, por muy buen desempeño que presente, es exacto). Es en este punto donde los tres modelos muestran un buen desempeño, y pueden ser útiles para predecir la potencia.

De acuerdo a los resultados obtenidos, los modelos predictivos multivariados son buenos en el modelamiento del molino SAG 1011, pero son muy complejos para ser utilizados y requieren de software estadísticos para ser implementados junto con mucho tiempo y demanda computacional. En contraste con el modelo propuesto por Austin, que a pesar de tener un error MAPE del 3,04% (mayor a los 2 creados), este no es significativamente mayor y resulta fácil y rápido de utilizar, donde además incluye factores fenomenológicos.

Finalmente, se propone tomar los datos obtenidos de nivel de llenado total en el solver del modelo de Austin (sección 4.4) y con ellos generar un modelo predictivo de nivel de llenado total con entradas de tratamiento (tph), velocidad del molino (rpm) y dureza (kWh/m^3). Con los supuestos resultados obtenidos estos pueden ocuparse de input para el modelo de Austin.

7 Bibliografía

- [1] **Amestica, R.**, “*Nuevo Modelos de Molienda Semiautógena para Simulación y Control*”, Tesis y Memoria de Título, Magister en Ciencias de la Ingeniería Mención Electricidad e Ingeniero Civil Electricista, Mayo 1995, Departamento de Ingeniería Eléctrica, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile.
- [2] **Amestica, R., Gonzalez, G.D., Menacho, J., Barria, J.**, “*A mechanistic state equation model for semiatogenous mill*”, *International Journal of Mineral Processing*, 44-45 (1996) 349-360.
- [3] **Barahona, C.**, *Descripción y Análisis de la Molienda Semiautógena*. Compañía Minera Disputada de las Condes S.A., Marzo 1991
- [4] **Breiman, L.** *Random Forests. Machine Learning*, Netherlands: Robert E. Schapire, 2001
- [5] **Benedetti, F., Balgún, A. & Yanguas, M.**, *Predicciones de Series de Tiempo con Redes Neuronales: Trading usando el Índice Merval.*, 2001
- [6] **Chin-Teng L., George Lee, C.S.**, “*Neural Fuzzy System: A Neuro-Fuzzy Synergism to Intelligent Systems*”, First Edition, April 1996
- [7] **Devijver, P. A., and J. Kittler**, *Pattern Recognition: A Statistical Approach*, Prentice-Hall, Londres, 1982
- [8] **Doll, A.** *An update data set for SAG mill power model calibration*. En: International Mineral Process Congress (18: 2016: Quebec, Canada). Paper 123. Logas Lake, Canada.
- [9] **Ersoy, O. K. & Hong, D.**, *Parallel, self-organizing, hierarchical neural networks. Neural Networks. IEEE Transactions on*, 1(2), pp. 167-178., 1999
- [10] **Freund, Y., y Schapire, R.** *Experiments with a New Boosting Algorithm*. En treceava conferencia internacional sobre machine learning, 148-156.
- [11] **Hastie, T., Friedman, J., y Tibshirani, R.** *The Elements of Statistical Learning*, Nueva York, Estados Unidos, 2001
- [12] **Kaastra, I. & Boyd, M.**, *Designing a neural network for forecasting financial and economic time series. Neurocomputing*, 10(3), pp. 215-236., 1996

- [13] **Mayer, D.G., Butler, D.G.** Statistical validation. En: *Ecological Modelling*. 1993. no. 68, pp. 21-32
- [14] **Quezada, J.**, *Preparación Mecánica de Minerales, Capítulos 1, 3, 4, 6*, Ingeniería en Metalurgia Extractiva, Universidad Arturo Prat.
- [15] **Rousseeuw, P; Leroy, A.**, *Robust Regression and Outlier Detection (3rd ed.)*, John Wiley & Sons, 1996
- [16] **Rokach, L., Maimon, O.**, *Data mining with decision trees: theory and applications*, Second Edition, 2008. ISBN 978-981-4590-07-5
- [17] **Tedeschi L.O.**, *Assessment of the adequacy of mathematical models*, 2006.
- [18] **Valencia, M., Yáñez, C. & Sánchez, L.**, *Algoritmo Backpropagation para Redes Neuronales: Conceptos y Aplicaciones*. Instituto Politécnico Nacional. Centro de Investigación en Computación, 2006
- [19] **Willis, MJ.**, "Artificial neural networks in process engineering" *IEEE Proceedings D (Control Theory and Applications)*, vol. 138 pp. 256 – 266, 1991
- [20] **R Core Team.** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>. 2017
- [21] **Fritsch,F., Guenther,F.**, neuralnet: Training of Neural Networks. R package version 1.33. <https://CRAN.R-project.org/package=neuralnet>. 2016
- [22] **Liaw,A., Wiener,M.**, Classification and Regression by randomForest. *R News* 2(3), 18--22. 2002

8 ANEXOS

8.1 Programa R

El programa estadístico R es un lenguaje de programación que permite llevar a cabo procedimientos estadísticos desde los más elementales hasta los más complejos; gracias a esto puede ser utilizado en diversos campos como la minería de datos. En el momento de redactar el código del programa, se emplearon funciones implementadas en los distintos *packages* que ofrece R. Resulta importante mencionar la utilidad de las librerías más destacadas.

8.2 Generación modelo de redes neuronales en R

```
# Instalar y llamar paquetes necesarios, en este caso "neuralnet"
install.packages('neuralnet')

library(neuralnet)

#Definir directorio de trabajo
dir <- "C:/Users/panch/Desktop/DataX"
setwd(dir)

#Cargar y Analizar los datos
datos = read.csv(file = 'DataX.csv', sep = ';', dec = ',')
N = ncol(datos)

for(j in 1:ncol(datos)){
  print(paste0(j,class(datos[,j])))}

maxs      = apply(datos, 2, max)
mins      = apply(datos, 2, min)
promedio  = apply(datos, 2, mean)
desvSt    = apply(datos, 2, sd)
estadisticos = cbind(mins,maxs,promedio,desvSt)

#Escalar datos
datos_escalados = as.data.frame(scale(datos,center = mins, scale = maxs -
mins))
```

```

#Definir tamaño de la muestra

proc_muestra = 70

muestra = sample.int(n = nrow(datos_escalados), size =
nrow(datos_escalados)%/(100/proc_muestra))

#-----
#           Var Salida Potencia
#-----

#Definir tabla de entrenamiento y de testeo
trainT = datos_escalados[ muestra , ]
testT = datos_escalados[-muestra , ]

#Entrenamiento de la red neuronal
red2 = neuralnet(Potencia ~ Tratamiento + Velocidad + Sólidos +
Granulometría + Dureza + Nivel.de.Bolas,
                trainT, Data
                hidden = 2, #Numero de neuronas en la capa oculta
                rep=3) #Numero de repeticiones

plot(red2, rep = 'best') #grafico de red

#Testeo de la red utilizando data nueva
prediccion = compute(red2, testT[,c(1,2,3,4,5,6)])
Tabla_red2 = data.frame(real=testT$Potencia,
                        Predicted = prediccion$net.result)

#Reescalar los datos a su forma original
varT = "Potencia" #Variable de salida que consideramos en el entrenamiento
pred_origT <- prediccion$net.result*(maxs[varT]-mins[varT])+mins[varT]
test_origT <- testT[,c("Potencia")]*(maxs[varT]-mins[varT])+mins[varT]

```

```

# Calculamos el error cuadratico medio
MSE.red <- 1/length(test_origT) * sum((test_origT - pred_origT)^2)
raizMSE.red <- sqrt(MSE.red)

Tabla_esc = data.frame(Prediccion = pred_origT , Original = test_origT)

#Extraer datos en excel
write.csv(Tabla_esc, file = 'Tabla_esc4.0.csv')

#-----#
#-----Validacion Cruzada-----#
#-----#

cantidad = 10 #Cantidad de iteraciones o k
MSE.T = vector("numeric",cantidad)
porc_muestra = 70
for(i in 1:cantidad ){
  #Definir tamano de la muestra
  muestra = sample.int(n = nrow(datos_escalados), size =
nrow(datos_escalados)%/(100/porc_muestra))
  #Definir tabla de entrenamiento y de testeo
  trainT = datos_escalados[ muestra , ]
  testT = datos_escalados[-muestra , ]
  #Entrenar la red
  red2 = neuralnet(Potencia ~ Tratamiento + Velocidad + Sólidos +
Granulometría + Dureza + Nivel.de.Bolas,
                    trainT,
                    hidden = 2,
                    rep=3)

```



```
#Calcular prediccion para la red entrenada
prediccion = compute(red2, testT[,c(1,2,3,4,5,6)])

#Reescalar los datos a su forma original
varT      = "Potencia" #Variable de salida que consideramos en el
entrenamiento
pred_origT <- prediccion$net.result*(maxs[varT]-mins[varT])+mins[varT]
test_origT <- testT[,c("Potencia")]*(maxs[varT]-mins[varT])+mins[varT]

# Calculamos el error cuadratico medio
MSE.red <- 1/length(test_origT) * sum((test_origT - pred_origT)^2)
MSE.T[i] = MSE.red
}

#Error obtenido por validacion cruzada
sqrt(mean(MSE.T))
```



8.3 Generación de Modelos Random Forest en R.

```

#Instalar y llamar paquete random forest
install.packages('randomForest')
library(randomForest)
#Definir el directorio de trabajo
dir <- "C:/Users/panch/Desktop/Data Z"
setwd(dir)

#Carga y análisis de datos
datos = read.csv(file = 'DataZ.csv', sep = ';', dec = ',') #Carga de datos
N = ncol(datos) #Verificación de columnas
x = nrow(datos) #Verificación de filas
for(j in 1:N){
  print(paste0(j,class(datos[,j]))) } #Verificación tipo de dato

#Calcular max, min, promedio, desc estandar y luego tabularlos en la
variable 'estadisticos'
maxs      = apply(datos, 2, max)
mins      = apply(datos, 2, min)
promedio  = apply(datos, 2, mean)
desvSt    = apply(datos, 2, sd)
estadisticos = cbind(mins,maxs,promedio,desvSt)

#Escalar datos
datos_escalados = as.data.frame(scale(datos,center = mins, scale = maxs -
mins))

#Definir tamaño de la muestra
proc_muestra = 70
muestra      = sample.int(n      = nrow(datos_escalados),      size      =
nrow(datos_escalados)%/(100/proc_muestra))

```

```

#-----
#           Var Salida Potencia
#-----

#Definir tabla de entrenamiento y de testeo
trainT = datos_escalados[ muestra , ]
testT  = datos_escalados[-muestra , ]

#-----
#           Entrenamiento Modelo Random Forest
#-----

RF_Power = randomForest(x = trainT[,c(1,2,3,4,5,6)], #Variables de entrada
en una matriz o data frame

                        y = trainT[,c(7)], #Variable de salida
                        ntree= 500, #Cantidad de arboles en el bosque
                        importance = TRUE)

#Grafico error versus numero de arboles
plot(RF_Power)

pred_RF_Power = predict(object = RF_Power, newdata = testT, type =
"response")

#Reescalar los datos a su forma original
varT = "Potencia" #Variable de salida que consideramos en el entrenamiento
pred_origT <- pred_RF_Power*(maxs[varT]-mins[varT])+mins[varT]
test_origT <- testT[,c("Potencia")]*(maxs[varT]-mins[varT])+mins[varT]

# Calculamos el error cuadratico medio
MSE.red <- 1/length(test_origT) * sum((test_origT - pred_origT)^2)
raizMSE.red <- sqrt(MSE.red)

```

```

#Histograma de errores
errorRF <- pred_origT - test_origT

TablaRF_comparativa = data.frame(Prediccion = pred_origT , Original =
test_origT)

#Extraer datos en excel
write.csv(TablaRF_comparativa, file = 'TablaRF_comparativa.csv')

#Variables de importancia
varImpPlot(RF_Power)
importance(RF_Power)

#-----
# Validacion cruzada
#-----
cantidad = 10
MSE.T = vector("numeric",cantidad)
porc_muestra = 70
for(i in 1:cantidad ){
  #Definir tamano de la muestra
  muestra = sample.int(n = nrow(datos_escalados), size =
nrow(datos_escalados)%/(100/porc_muestra))

  #Definir tabla de entrenamiento y de testeo
  trainT = datos_escalados[ muestra , ]
  testT = datos_escalados[-muestra , ]

  #Entrenar el bosque
  RF_Trat = randomForest(x = trainT[,c(1,2,3,4,5,6)], #Variables de entrada
en una matriz o data frame
                          y = trainT[,c('Potencia')], #Variable de salida
                          ntree= 150) #Cantidad de arboles en el bosque

  #Calcular prediccion para el bosque entrenado
  pred_RF_Trat = predict(object = RF_Trat, newdata = testT, type =
"response")

```

```
#Reescalar los datos a su forma original

varT = "Potencia" #Variable de salida que consideramos en el
entrenamiento

pred_origT <- pred_RF_Trat*(maxs[varT]-mins[varT])+mins[varT]
test_origT <- testT[,c("Potencia")]*(maxs[varT]-mins[varT])+mins[varT]

# Calculamos el error cuadratico medio
MSE.red <- 1/length(test_origT) * sum((test_origT - pred_origT)^2)
MSE.T[i] = MSE.red
}

#Error obtenido por validacion cruzada
sqrt(mean(MSE.T))
```



UNIVERSIDAD DE CONCEPCIÓN – FACULTAD DE INGENIERÍA**Departamento de Ingeniería Metalúrgica**

Hoja Resumen Memoria de Título

Título: Diseño de modelo predictivo multivariable para molienda en Compañía Minera Doña Inés de Collahuasi
Nombre Memorista: Francisca Valentina Contreras Acchiardo

Modalidad		Profesor(es) Patrocinante(s)
Concepto		
Calificación		
Fecha	10 de Agosto del 2018	
Prof. ^a Eugenia Araneda H.		Prof. Fernando Betancourt Cerda
		Ingeniero Supervisor: Nelson Mora Navarrete
		Institución: Compañía Minera Doña Inés de Collahuasi

Comisión (Nombre y Firma)	
Prof. Daniel Sbarbaro	Prof. Roberto Fustos

Resumen
<p>El presente trabajo se desarrolló en las dependencias de la gerencia de geometalurgia, de CMDIC con el objetivo de desarrollar un modelo para predecir el comportamiento no lineal de la potencia de uno de los molinos SAG, teniendo como entradas las variables mineralógicas junto con las variable manipulables por el operador y como salida única la potencia.</p> <p>La idea surge debido a que la planta no cuenta con modelo de potencia en los molinos SAG, lo que conlleva a una variabilidad en la operación, error en las planificaciones y estimaciones de tratamiento.</p> <p>Se proponen tres tipos de modelos predictivos a comparar: Redes Neuronales y Bosques Aleatorios y el establecido por Austin. Para la realización de los modelos se utilizó data real recopilada de planta. Previamente al entrenamiento de los modelos la data cruda se sometió a un tratamiento para eliminar outliers.</p> <p>El criterio de comparación de los modelos fue la Raíz del Error Cuadrático Medio y el Error Porcentual Absoluto Medio.</p> <p>Como resultado se tiene que el método Random Forest con 150 árboles de decisión tiene un error de 0,76%, el de Redes Neuronales 1,40% y Austin 3,04%.</p> <p>Finalmente, como la potencia depende del nivel de llenado total del molino, se propone desarrollar un modelo predictivo de nivel de llenado total, lo que contribuye de mejor manera con el modelamiento de la potencia del molino.</p>