
**Predicción de la resistencia mecánica a 28 días del cemento Concretero a
través de regresión lineal múltiple en un molino vertical con base en las
propiedades físicas y químicas del cemento**



Presentado por

Miller Albeiro Puentes Lozano

LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Fundación Universitaria Los Libertadores

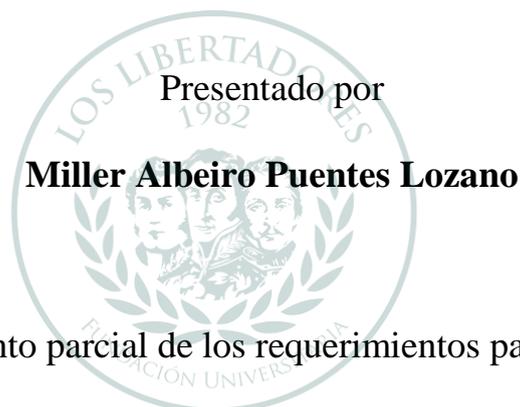
Facultad de Ingeniería y Ciencias Básicas

Especialización en Estadística Aplicada

Bogotá D.C, Colombia

2018

Predicción de la resistencia mecánica a 28 días del cemento Concretero a través de regresión lineal múltiple en un molino vertical con base en las propiedades físicas y químicas del cemento



Presentado por

Miller Albeiro Puentes Lozano

En cumplimiento parcial de los requerimientos para optar al título

Especialista en Estadística Aplicada

LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA
Dirigida por

Jessica Ximena Moreno

Fundación Universitaria Los Libertadores

Facultad de Ingeniería y Ciencias Básicas

Especialización en Estadística Aplicada

Bogotá D.C, Colombia

2018

Notas de aceptación



LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Firma del presidente del jurado

Firma del jurado

Firma del jurado

Bogotá D.C, Diciembre de 2018



LOS LIBERTADORES

FUNDACIÓN UNIVERSITARIA

Las directivas de la Fundación Universitaria Los Libertadores, los jurados calificadores y el cuerpo docente no son responsables por los criterios e ideas expuestas en el presente documento. Estos corresponden únicamente a los autores y a los resultados de su trabajo.

Dedicatoria

A Dios.

Por darme la oportunidad y la salud para alcanzar las metas que me he propuesto y que intento llevar de la mano con su propósito para mí.

A mi novia Nathalia.

Por su paciencia, amor y comprensión; pero sobre todo por mostrarme el amor de Dios en todo lo que hace.

A mi padre Albeiro.

Por haber formado en mí la persona responsable y persistente que termina lo que inicia a pesar de las adversidades.

A mi madre Jacqueline.

Por su ejemplo de dedicación y por expresar su amor incondicional en todo momento.

A mis hermanos Jacquelin y Kevin.

Porque me inspiran a seguir creciendo como profesional y a nivel personal para darles un mejor ejemplo.

Agradecimientos

Gracias a la compañía cementera donde me desempeñé, gracias por haberme permitido sacar a delante este trabajo con la información de su proceso y formarme para aplicar los conocimientos en la compañía, gracias a todas las personas que fueron partícipes de este proyecto, ya sea de manera directa o indirecta, gracias a la universidad Fundación Universitaria Los libertadores por permitirme ser parte de sus egresados.

LOS LIBERTADORES
FUNDACIÓN UNIVERSITARIA

Predicción de la resistencia mecánica a 28 días del cemento Concretero a través de regresión lineal múltiple en un molino vertical con base en las propiedades físicas y químicas del cemento

Resumen

Este trabajo parte de la necesidad de predecir la resistencia del cemento Concretero a 28 días en un molino vertical al examinar la relación existente entre las variables químicas y físicas medibles en él, con el objetivo de tomar decisiones que impacten positivamente en los costos de producción como el aumento en las adiciones de materiales inertes que no aportan a la resistencia mecánica del cemento.

La fundamentación teórica se realizó desde el campo de estudio de la industria cementera y la estadística tanto descriptiva como multivariante. Con esto, se busca construir un modelo que posibilite innovación dentro de los procesos de producción de cemento, así como la disminución de los impactos negativos que tienen consecuencias en el medio ambiente y en la rentabilidad del negocio.

Desde esta perspectiva, y teniendo en cuenta los avances del Big Data. Se crea la necesidad de tener herramientas que ayuden a tomar decisiones basados en métodos estadísticos realizando análisis de los datos que se registran desde el área de calidad. Los ingenieros que trabajan en plantas de producción enfrentan día a día la toma de decisiones y con regularidad no se cuenta con el tiempo necesario para evaluar todas las posibilidades e impactos de la decisión a tomar, sin embargo, la evaluación de modelos predictivos facilita comprender como afecta la toma y ejecución de determinada decisión.

En este sentido y para avanzar en la resolución del problema, se planteó un interrogante para conocer el comportamiento de la resistencia mecánica del cemento Concretero en función de propiedades medibles en el día a día, las cuales nos permitieron operativizar el desarrollo del trabajo y diseñar un modelo experimental y predictivo para obtener los resultados que se muestran a lo largo del documento. Finalmente, encontrará la contribución del trabajo y propuesta final, para terminar con relación bibliográfica utilizada en el trabajo de investigación.

Palabras clave: regresión lineal multivariada, C3S, C2S, C3A, C4AF, Blaine, resistencia mecánica, cemento Concretero, molienda vertical.

Tabla de contenido

ÍNDICE DE TABLAS	10
ÍNDICE DE FIGURAS	11
ÍNDICE DE ECUACIONES	12
1. INTRODUCCIÓN	13
2. PLANTEAMIENTO DEL PROBLEMA	14
2.1 OBJETIVOS	16
2.1.1 OBJETIVO GENERAL	16
2.1.2 OBJETIVOS ESPECÍFICOS	16
2.2 JUSTIFICACIÓN	17
3. MARCO TEÓRICO	19
3.1 MOLIENDA DE CEMENTO	19
3.1.1 MOLIENDA HORIZONTAL O DE BOLAS	19
3.1.2 MOLIENDA VERTICAL	20
3.1.3 PROPIEDADES FÍSICAS Y MECÁNICAS DEL CEMENTO	20
3.1.4 PROPIEDADES QUÍMICAS DEL CEMENTO	21
3.2 REGRESIÓN LINEAL MÚLTIPLE	25
3.2.1 IDENTIFICACIÓN DE DATOS ATÍPICOS (MÉTODO DE MAHALANOBI)	26
3.2.2 ESTADÍSTICA DESCRIPTIVA	28
3.2.3 EL MODELO DE REGRESIÓN LINEAL MÚLTIPLE Y SU ESTIMACIÓN POR MÍNIMOS CUADRADOS.	30
3.2.4 SUPUESTOS BÁSICOS DEL MODELO	31
4. MARCO METODOLÓGICO	33
4.1 FASE I. IMPLEMENTACIÓN DE LAS ECUACIONES DE BOGUÉ EN LOS DATOS QUÍMICOS.	34
4.2 FASE II. CARACTERIZACIÓN DEL PROCESO Y ANÁLISIS DE VARIABLES.	34
4.3 FASE III. DETECCIÓN DE CASOS ATÍPICOS.	34
4.4 FASE IV. SELECCIÓN DE VARIABLES.	35
4.5 FASE V. CREACIÓN DEL MODELO DE REGRESIÓN.	35
4.6 FASE VI. VALIDACIÓN DEL MODELO.	35
5. ANÁLISIS Y RESULTADOS	36
5.1 CALCULAR LA COMPOSICIÓN QUÍMICA DEL CEMENTO CONCRETERO POR EL MÉTODO DE BOGUÉ. (C3S, C2S, C4AF, C3A)	36
5.2 REALIZAR ANÁLISIS DESCRIPTIVO DE LAS VARIABLES QUE INFLUYEN EN LA RESISTENCIA A 28 DÍAS EN EL CEMENTO CONCRETERO.	37
5.3 REALIZAR ANÁLISIS MULTIVARIADO PARA ELEGIR LAS VARIABLES DE MAYOR INFLUENCIA EN LA RESISTENCIA A 28 DÍAS.	48
5.4 VALIDAR EL MODELO DE REGRESIÓN LINEAL PLANTEADO.	55
5.4.1 INDEPENDENCIA DE LOS RESIDUOS	55
5.4.2 NORMALIDAD DE LOS RESIDUOS	57
5.4.3 IGUALDAD DE VARIANZAS	58
5.4.4 MULTICOLINEALIDAD	60

6. CONCLUSIONES Y RECOMENDACIONES	62
BIBLIOGRAFÍA	66
ANEXOS	69

Índice de tablas

Tabla N. 1 Los 4 principales compuestos del Cemento	22
Tabla N. 2 Ecuaciones de Bogue aplicadas al cemento	37
Tabla N. 3 Variables del cemento que influyen en la resistencia a 28 días	37
Tabla N. 4 Estadísticos de las primeras 5 variables a analizar	39
Tabla N. 5 Estadísticos de las siguientes 5 variables por analizar	39

Índice de figuras

Figura N.1. Datos en RStudio de la distancia de Mahalanobi	38
Figura N.2. Diagrama en RStudio % de partículas menor a $3\mu\text{m}$	40
Figura N.3. Diagrama en RStudio % de partículas entre $3\mu\text{m}$ y $30\mu\text{m}$	40
Figura N.4. Diagrama en RStudio % de partículas mayor $30\mu\text{m}$	41
Figura N.5. Diagrama en RStudio Blaine (cm^2/g).....	42
Figura N.6. Diagrama en RStudio retenido en malla 325	42
Figura N.7. Diagrama en RStudio Resistencia a 28 días	43
Figura N.8. Diagrama en RStudio Contenido de Alita C3S	44
Figura N.9. Diagrama en RStudio contenido de Belita C2S.....	45
Figura N.10. Diagrama en RStudio contenido de Celita C3A	45
Figura N.11. Diagrama en RStudio contenido de Felita C4AF	46
Figura N.12. Diagrama en RStudio pérdida ígnea	47
Figura N.13. Diagrama en RStudio de Cal libre	47
Figura N.14. Diagrama en RStudio relación Sodio-Potasio EA	48
Figura N.15. Datos en RStudio Matriz de covarianza $d_3\mu\text{m}$, $d_{3-30\mu\text{m}}$, $d_{30\mu\text{m}}$, Blaine, Ret325	49
Figura N.16. Datos en RStudio Matriz de covarianza R28, C3S, C2S, C3A, C4AF.....	49
Figura N.17. Datos en RStudio Matriz de covarianza Pérdida ígnea, Cal libre, EA	49
Figura N.18. Datos en RStudio Matriz de correlación $d_3\mu\text{m}$, $d_{3-30\mu\text{m}}$, $d_{30\mu\text{m}}$, Blaine, Ret325, R28	50
Figura N.19. Datos en RStudio Matriz de correlación C3S, C2S, C3A, C4AF, Pérdida ígnea, cal libre.....	50
Figura N.20. Datos en RStudio Matriz de correlación EA	50
Figura N.21. Diagrama de dispersión $d_3\mu\text{m}$, $d_{3-30\mu\text{m}}$, $d_{30\mu\text{m}}$, Blaine, Ret325, R28	51
Figura N.22. Diagrama de dispersión C2S, C3S, C3A, C4AF, Pérdida ígnea, cal libre, EA .	51
Figura N.23. Datos en RStudio modelo de regresión lineal.....	52
Figura N.24. Datos en RStudio, segundo modelo de regresión lineal	53
Figura N.25. Datos en RStudio segundo modelo de regresión lineal sin la variable Pérdida ígnea.....	54
Figura N.26. Datos en RStudio Modelo de regresión lineal sin C3S.....	55
Figura N.27. Resultado del modelo en RStudio.....	56
Figura N.28. Diagrama test Durbin-Watson en RStudio de Blaine, retenido 325, relación entre potasio y sodio.	56
Figura N.29. Diagrama test Durbin-Watson en RStudio de C2S, C3A, C4AF	57
Figura N.30. Resultado en RStudio normalidad de los residuos.....	57
Figura N.31. Diagrama en RStudio de los residuos del modelo	58
Figura N.32. Gráfico igualdad de varianzas.....	59
Figura N.33. Resultados en RStudio para igualdad de varianzas	59
Figura N.34. Resultados RStudio multicolinealidad.....	60
Figura N.35. Gráfico de dispersión multicolinealidad Blaine, Ret 325, C3S	61
Figura N.36. Gráfico de dispersión multicolinealidad C3A, C4AF, EA	61

Índice de ecuaciones

Ecuación N.1. Ecuación de Bogué del C3S	23
Ecuación N.2. Ecuación de Bogué del C2S	23
Ecuación N.3. Ecuación de Bogué del C3A.....	24
Ecuación N.4. Ecuación de Bogué para el C4AF	24
Ecuación N.5. Ecuación de la distancia de Mahalanobis	27
Ecuación N.6. Ecuación de la matriz de covarianza muestral.....	28
Ecuación N.7. Ecuación matemática del promedio	29
Ecuación N.8. Modelo de regresión lineal múltiple MCO	31
Ecuación N.9. Ecuación del supuesto de independencia entre los residuos.....	31
Ecuación N.10. Ecuación Modelo de regresión lineal elegido.....	55

1. Introducción

En la actualidad sobre la producción sostenible se trabajan diferentes enfoques que posibilitan e impulsan nuevos modelos de productividad que permitan mejores ganancias dentro de las compañías y a su vez contribuyan o sean amigables con el medio ambiente. La industria cementera no es ajena al desarrollo ni a la implementación de herramientas que ayuden a ser sostenibles y de mejores rentabilidades para los accionistas. A partir de esto se plantea desde la estadística un trabajo académico e investigativo, que tiene como finalidad predecir una variable física importante en uno de sus productos terminados.

La compañía de cementos en donde se realiza este trabajo investigativo produce tres tipos de cemento a saber: Cemento de Uso General, Cemento Estructural y Cemento Concretero, éste último como punto de interés particular para esta investigación puesto que es el más costoso de producir debido a la alta exigencia en las resistencias a 28 días que se tiene para satisfacer a los clientes que compran el producto. Además de tener tres tipos de cementos en su línea de producción, cuenta con 2 tipos de molienda que tienen principios de operación diferentes entre sí y afecta el desempeño en cuanto a resistencias del cemento producido. Sin embargo, el equipo que se destina para una mayor producción anual del cemento en cuestión en la planta, es el molino vertical de allí el interés para este proyecto.

Por lo tanto, en la compañía de Cementos donde se realiza esta investigación las pretensiones van orientadas a la conformación de un modelo de regresión estadístico que permita predecir la resistencia del cemento Concretero a 28 días en función de las variables químicas y físicas que se miden regularmente en el día a día de la producción en el molino vertical. Lo anterior con el fin de enfocar los esfuerzos en conocer el comportamiento del cemento producido y realizar ajustes en las dosificaciones de la materia prima o desde procesos para cumplir con

las exigencias de calidad.

De este modo, en la presente investigación se parte de la siguiente pregunta de investigación ¿Cuál es la relación existente entre las variables químicas y físicas y la resistencia del cemento Concretero a 28 días en un molino vertical?, a partir de esto se establece el objetivo general de diseñar un modelo de regresión lineal para predecir la resistencia mecánica a 28 días del cemento Concretero en un molino vertical con base en las propiedades químicas y físicas del cemento, para dar cumplimiento a este aspecto, se realiza una metodología investigativa que pretende tomar partido de datos cuantitativos para poder de esta manera comprender en mayor medida el planteamiento ya mencionado.

En el segundo capítulo del trabajo se establece el planteamiento del problema, los objetivos y la justificación. En el tercer capítulo se presenta el marco teórico que fundamenta el trabajo, donde se realiza la distinción de los aspectos teóricos en relación a la molienda de cemento y los modelos de regresión lineal multivariante, en el cuarto capítulo, se describe la propuesta metodológica para resolver el problema y los objetivos. En el capítulo quinto se presentan los resultados dando cumplimiento a los objetivos específicos planteados a la luz de la estadística descriptiva y multivariante, registrando y validando el modelo de regresión obtenido a la luz de los supuestos de un modelo de regresión lineal multivariante.

Finalmente, se presentan las conclusiones, las recomendaciones y los referentes bibliográficos, para terminar con los anexos.

2. Planteamiento del problema

Esta investigación se llevó a cabo en una compañía cementera colombiana con más de 70 años de experiencia en el mercado nacional y con el quinto puesto en producción de cemento

en América Latina. La planta para la que se desarrolla el modelo tiene más de 30 años operando y produciendo entre su portafolio de productos cemento de uso general, cemento de uso estructural y cemento concretero.

Esta compañía tiene como uno de sus pilares fundamentales el compromiso de cumplir con la calidad de sus productos satisfaciendo los requerimientos de los clientes. Ya que el impacto que tiene como gestor de cambio social es grande y de mucha responsabilidad, se ciñe a la normatividad nacional que regula la calidad del cemento mediante las normas técnicas colombianas (NTC) 121 y 321.

Las materias primas para la producción de cemento no son homogéneas, por ello se hace difícil predecir con certeza o con algún grado de confiabilidad el comportamiento del producto terminado sin nada más que por la experiencia del personal humano involucrado en la producción y la literatura que se ha generado a lo largo del desarrollo de esta industria. Por lo anterior se pone de manifiesto la necesidad de conocer un modelo que permita conocer el comportamiento en el tiempo de sus productos producidos con base en sus propiedades químicas y físicas que permita determinar qué resistencia puede ofrecer el cemento Concretero a sus clientes, de acuerdo con su capacidad de proceso, con el objetivo de poder producir de manera sustentable, un poco más rentable y cumpliendo las normativas nacionales.

La planta dentro de su organización cuenta con el área de control de calidad, ésta a su vez se encuentra conformada por dos secciones de análisis; laboratorio de análisis químico y laboratorio de análisis físico, debido a que surge la necesidad de mantener condiciones específicas para cada tipo de análisis, con el fin de apegarse a lo establecido en las normas

técnicas colombianas.

El área de control de calidad de la compañía, cuenta con una extensa base de datos que se ha recopilado a lo largo de los años de las mediciones químicas y físicas de sus materias primas, productos intermedios y productos terminados, esto con el fin de tener controlado todo el proceso de producción para evitar cualquier desviación en la fabricación del cemento y así cumplir con los estándares de calidad.

Para la producción del cemento Concretero, la exigencia más importante por NTC y por requerimientos de los clientes es la resistencia a 28 días, esta propiedad física es fundamental ya que dicta el comportamiento futuro de las estructuras realizadas.

En el desarrollo del trabajo se abordarán las propiedades tanto químicas como físicas que influyen en el desempeño de esta resistencia.

2.1 Objetivos

2.1.1 Objetivo General

Diseñar un modelo de regresión lineal para predecir la resistencia mecánica a 28 días del cemento Concretero en un molino vertical con base en las propiedades químicas y físicas del cemento.

2.1.2 Objetivos Específicos

- Calcular la composición química del cemento Concretero por el método de Bogue.
(C_3S , C_2S , C_4AF , C_3A)
- Realizar análisis descriptivo de las variables que influyen en la resistencia a 28 días en el cemento Concretero.
- Realizar análisis de la regresión lineal para elegir las variables de mayor influencia en

la resistencia a 28 días.

- Validar el modelo de regresión lineal planteado.

2.2 Justificación

Según la Norma Técnica Colombiana NTC 121 el cemento es un producto que se obtiene por la pulverización del Clinker Portland con la adición de una o más formas de sulfato de calcio. Se admite la adición de otros productos siempre que su inclusión no afecte las propiedades del cemento resultante. Todos los productos adicionales deben ser pulverizados conjuntamente con el Clinker.

Cuando el cemento portland es mezclado con el agua, se obtiene un producto de características plásticas con propiedades adherentes que fragua y endurece a causa de las reacciones de hidratación de sus constituyentes, dando lugar a productos mecánicamente resistentes. Predecir dicha resistencia ha sido tema de estudio por varios años en diferentes países alrededor del mundo.

Lo anterior, se convierte en un tema de interés ya que permite adelantarse a tomar decisiones que disminuyan los costos de producción generando primeramente una mayor rentabilidad en el negocio y en segunda instancia un impacto positivo en el cuidado del ambiente. Países como Guatemala, en donde se llevó a cabo una investigación por Valle García, en 2018 titulada “Modelo de regresión lineal multivariado aplicado a la relación entre variables químicas del Clinker y variables físicas del cemento” en el que se desarrolla un modelo de regresión lineal para las variables que influyen en el desempeño de la resistencia del cemento a 28 días teniendo en cuenta la calidad del Clinker, da pie para seguir investigando y ahondando en el tema de predecir la resistencia del cemento a 28 días para dar cumplimiento

a las NTC que rigen la producción del cemento en nuestro territorio colombiano.

Justificando la investigación a presentar, a continuación, se describen dos ámbitos sumamente importantes por lo cual se decide llevar dicha investigación:

1. Desde el ámbito ambiental: Se sabe que la industria cementera es una de las grandes generadoras de CO₂ como lo enuncian Otero, Ramirez, Bukasa, & Seke, 2016, en el artículo “Efecto de la adición de caliza a Clinker portland Angolano en el comportamiento mecánico de morteros” en el que se le atribuye la responsabilidad de generar el 7 % de las emisiones de CO₂ a la atmósfera. También menciona que aproximadamente el 40% de las emisiones de CO₂ corresponden a la energía empleada para operar y el 60% restante surge como descomposición térmica de la materia prima para formar el Clinker.
2. Desde el ámbito económico: En la industria cementera como en cualquier otra, existe competencia por ser atractivo financieramente para los inversionistas. Esto hace que cada día se piense en reemplazar las materias primas de mayor valor económico por materiales que desempeñen iguales o mejores comportamientos en el desempeño del cemento o en el peor de los casos no influyan en su desempeño. En la molienda de cemento Concreteero se tiene una mezcla de 6% de yeso, 12% de Caliza de adición y 82% de Clinker. Siendo este último el de mayor valor total sobre el cemento producido. Al tener una buena estimación de la resistencia del cemento a 28 días, se puede variar la receta aumentando el porcentaje de Caliza de adición y bajando el contenido de Clinker, disminuyendo así el costo de producción.

3. Marco teórico

3.1 Molienda de Cemento

El proceso de fabricación del cemento consiste en tomar materiales calcáreos y arcillosos de la naturaleza, transformarlos en polvo con determinada composición química, homogeneizarlos y someterlos a un proceso de cocción a temperaturas del orden de 1500°C, para obtener el Clinker.

El cemento como tal es el resultado de la molienda fina del Clinker con un material que regule su fraguado, como el yeso, y otras adiciones como calizas, puzolanas, escorias, cenizas, etc; que hacen parte de las necesidades del usuario final y de los compromisos de calidad que se adquieran.

Dentro de la industria cementera se cuenta con dos tipos de molienda de cemento, a saber:

3.1.1 Molienda horizontal o de bolas

Es el molino de cemento más utilizado ya que se adapta a diversas aplicaciones en las plantas, y tiene simplicidad de funcionamiento y un fácil mantenimiento. El molido toma su nombre por un cilindro hueco horizontal giratorio que en su interior contiene cuerpos moledores metálicos de diferentes tamaños. Los cuerpos moledores más grandes son los encargados de moler las partículas más grandes y por otra parte los cuerpos moledores más pequeños se encargan de las partículas más pequeñas. El material que se va a moler se alimenta por la entrada del molino, después de esto el material junto con los cuerpos moledores son llevados por el casco giratorio del molino. Mientras el cilindro da vueltas las partículas del material (Clinker) se rompen al ser impactadas por los cuerpos moledores y el molino. El molino horizontal o de bolas tiene el consumo energético más alto que el de otro tipo de molino

(vertical), tiene un medio especial y específico para el sistema de enfriamiento y ventilación, y, además requiere una maquinaria adicional por el funcionamiento en circuito cerrado para el transporte y clasificación del producto del molino (CPI, 2003).

3.1.2 Molienda vertical

El molino vertical en términos de rentabilidad máxima de toda la planta es el único que permite una integración total en el sistema de regulación. Los molinos verticales nacieron inicialmente para moler carbón, cuya abrasividad es muy superior a la del crudo de cemento, algunas otras aplicaciones de este molino son para Clinker, yeso y puzolanas. El consumo energético del molino vertical es mucho menor que el del molino horizontal, esto se debe a que produce menos finas las partículas, y que estas partículas cuando están molidas son arrastradas por la corriente de aire a través de un separador. El mecanismo de un molino vertical consta de un juego de rodillos sobre una mesa giratoria, la recogida del producto se efectúa por elevador que transporta el material molido al molino de cemento una vez que cae por el borde de la mesa (Hidalgo de Cisneros Alonso, 2003).

3.1.3 Propiedades físicas y mecánicas del cemento

- **Superficie específica “Blaine”**

Las partículas de cemento, debido a su pequeño tamaño, no pueden caracterizarse por medio de tamices; por esta razón, se necesitan otros métodos para medir el tamaño de partícula.

El método Blaine para cemento se utiliza para la determinación de la superficie específica del Cemento. El ensayo se lleva a cabo en el permeabilímetro de aire, y el resultado se compara con un estándar (Otro Cemento de superficie específica conocida como patrón) (Cebrian & Pisonero, 1971).

El método del aparato Blaine como lo describe la norma E-302-07 (2008) “Consiste en hacer pasar una cantidad determinada de aire a través de una capa de cemento de porosidad definida. La cantidad y el tamaño de los poros existentes en dicha capa, son función del tamaño de las partículas y determinan el gasto del aire a través de la capa”.

- **Distribución granulométrica**

La granulometría o gradación como lo define Gutierrez De Lopez, 2003 se refiere al tamaño de las partículas y al porcentaje o distribución de las mismas en una cantidad de cemento. Esta propiedad se determina mediante el análisis granulométrico que consiste en hacer pasar una determinada cantidad del cemento a través de una serie de tamices standard, dispuestos de mayor a menor.

- **Resistencia mecánica**

La resistencia mecánica del cemento es la capacidad del cemento para resistir esfuerzos a la compresión. Al combinar el cemento con agua se da lugar a pastas que se endurecen y hacen cohesión entre sus partículas y, cuya porosidad disminuye a medida que las resistencias mecánicas van creciendo en el tiempo (Neville, 2013).

3.1.4 Propiedades químicas del cemento

La identificación de los principales compuestos del Clinker de Cemento se basa en gran medida en las ecuaciones de Bogue y, por lo tanto, se les conocen como "Compuestos de Bogue". Los cuatro compuestos generalmente considerados son:

Tabla N. 1*Los 4 principales compuestos del Cemento*

Nombre del compuesto	Ecuación química	Símbolo
Silicato tricálcico	$3 \text{ CaO} \cdot \text{SiO}_2$	C_3S
Silicato dicálcico	$2 \text{ CaO} \cdot \text{SiO}_2$	C_2S
Aluminato tricálcico	$3 \text{ CaO} \cdot \text{Al}_2\text{O}_3$	C_3A
Ferro aluminato tetracálcico	$4 \text{ CaO} \cdot \text{Al}_2\text{O}_3 \cdot \text{Fe}_2\text{O}_3$	C_4AF

En este trabajo, se aplican las ecuaciones de Bogue no al Clinker con el que se produce el cemento, sino directamente al cemento Concretero producido. Con esto, se busca incluir los compuestos de mayor importancia en la industria cementera en el análisis de regresión.

A continuación, se describe cada uno de estos compuestos y la ecuación planteada por Bogue para determinar sus porcentajes presentes en el cemento:

- **Silicato tricálcico (C3S)**

También llamado Alita, este componente del cemento controla las resistencias iniciales y el endurecimiento, por esta razón presenta un leve calor de hidratación. Este es el constituyente del cemento que libera la mayor cantidad de CaO (óxido de calcio reactivo). La composición química de este compuesto es la combinación de dos óxidos 3CaO y SiO_2 . Si un cemento contiene más del 60% de Silicato tricálcico, el resultado será que se endurecerá bastante rápido y con altas resistencias iniciales, y como resultado de esto el cemento se desmoldará más rápido en la obra. El C_3S puro mezclado con agua también contribuye al fraguado inicial (Cortes Gomez & Perilla Sastoque, 2014).

Según Bogue el porcentaje de C_3S presente en el Clinker se puede determinar como

(Stutzman, Heckert, Tebbe, & Leigh, n.d.):

Ecuación N.1. *Ecuación de Bogue del C3S*

$$C_3S = 4,07(CaO) - 7,60(SiO_2) - 6,72(Al_2O_3) - 1,43(Fe_2O_3) - 2,85(SO_3)$$

- **Silicato Bicálcico (C2S):**

También llamado Belita. El C₂S se endurece en forma más gradual que el C₃S. “La composición del óxido es 2CaO.SiO₂, este compuesto libera bajo calor de hidratación 60 cal/g, con una velocidad intrínseca de reacción de 1cal/g/h siendo lenta su hidratación y endurecimiento, por lo que el aumento en la resistencia se adquiere después de la primera semana aproximadamente” (Cortes Gomez & Perilla Sastoque, 2014).

Según Bogue el porcentaje de C₂S presente en el Clinker se puede determinar cómo (Stutzman et al., n.d.):

Ecuación N.2. *Ecuación de Bogue del C2S*

$$C_2S = 2,87(SiO_2) - 0,754(C3S)$$

- **Aluminato tricálcico (C3A):**

También llamado Celita. Este compuesto agrietado se encuentra en forma de láminas planas. El aluminato tricálcico reacciona muy rápidamente con el agua, generando mucho calor, no contribuye mucho con la resistencia inicial. Ciertos compuestos al reaccionar con agua reaccionan de manera violenta e inmediatamente lleva a un endurecimiento de la pasta, éste es el caso del C₃A. Para retrasar este proceso del C₃A se debe emplear una cantidad propicia de yeso que se agrega al Clinker. Un aumento de la finura del C₃S eleva la cantidad de yeso que se requiere para retardar el efecto de endurecimiento instantáneo (Cortes Gomez & Perilla

Sastoque, 2014).

El C₃A facilita la combinación de cal y sílice; debido a esto es de vital importancia en la fabricación del cemento (Neville, 2013). Según Bogue el porcentaje de C₃A presente en el Clinker se puede determinar cómo (Stutzman et al., n.d.):

Ecuación N.3. Ecuación de Bogue del C₃A

$$C_3A = 2,65(Al_2O_3) - 1,69(Fe_2O_3)$$

- **Ferroaluminaro tetracálcico (C₄AF):**

También llamado Ferrita. “Se encuentra del 5 – 15% en el cemento, la composición del óxido es 4CaO.Fe₂O₃.Al₂O₃, este compuesto actúa como fundente, y de esta forma reduce la temperatura de cocción del Clinker. A mayor cantidad C₄AF en el cemento lo hace más vulnerable al ataque químico por sulfatos” (Cortes Gomez & Perilla Sastoque, 2014). Este compuesto contribuye poco a la resistencia temprana del cemento ya que se caracteriza por una velocidad de reacción lenta con un moderado calor de hidratación.

Según Bogue el porcentaje de C₄AF presente en el Clinker se puede determinar cómo (Stutzman et al., n.d.):

Ecuación N.4. Ecuación de Bogue para el C₄AF

$$C_4AF = 3,04(Fe_2O_3)$$

- **Relación Sodio – Potasio (EA):**

También llamado Álcalis (Na₂O + K₂O), esta es una reacción que ocurre entre el sodio y el

potasio del cemento con los agregados silíceos. Esta reacción interfiere en la rapidez con que el cemento adquiere resistencia. La reacción ayuda a reaccionar con algunos agregados del cemento lo que ayuda a la desintegración del mismo. Los álcalis afectan en la reacción del C_3S ayudando a que sea mayor la reactividad con concentración y solubilidad más rápida, o, por otro lado, puede influir en la reducción de la fluidez (Cortes Gomez & Perilla Sastoque, 2014).

- **Perdida por calcinación:**

También llamada Cal libre u óxido de calcio (CaO), son componentes no deseados que indican una cocción insuficiente del Clinker, o un enfriamiento ineficiente que conlleva al rompimiento de la Celita. El óxido de calcio cuando está en niveles altos produce la expansión del cemento después del fraguado final. El CaO en su estado natural reacciona instantáneamente con el agua, si esta es calcinada, su área superficial se reduce. “La cal libre afecta principalmente la resistencia del cemento en especial por la falta de solidez, estabilidad volumétrica y posibles agrietamientos, de acuerdo a investigaciones el problema es que a los dos días la resistencia disminuye, y esto toma mayor importancia a los 28 días en adelante. Para evitar esto la cal libre en el cemento portland debe ser limitada para evitar la expansión, que puede ocurrir durante el endurecimiento”(Cortes Gomez & Perilla Sastoque, 2014).

3.2 Regresión lineal múltiple

La regresión lineal múltiple como lo menciona Montero Granados (2016), en su artículo “Modelos de regresión múltiple” trata de ajustar modelos lineales o linealizables entre más de una variable explicativa (variable dependiente) y más de una variable independiente a diferencia de la regresión lineal simple que analiza la influencia de una variable explicativa (X) en los valores que toma otra variable denominada dependiente (Y). La regresión lineal

múltiple ofrece la ventaja de utilizar más información (variables) en la construcción del modelo y, por esta razón, realizar estimaciones más concretas.

3.2.1 Identificación de datos atípicos (método de Mahalanobi)

Para las organizaciones es de gran importancia realizar el análisis de calidad de los datos en sus procesos de producción o/y de servicios, ya que los datos fuera de los estándares establecidos pueden conducir a decisiones erróneas, productos con problemas, desatenciones a los clientes, etc. Como resultado de un mal análisis o de evadir datos las empresas llegan a una pérdida de dinero, tiempo y credibilidad. Por esta razón es importante identificar los datos atípicos (outliers) de los procesos en cualquier organización. Hawkins (1980) define Outliers como “una observación que se desvía mucho de otras observaciones y despierta sospechas de ser generada por un mecanismo diferente”; aunque estos valores pueden aparentar ser inválidos pueden ser correctos y viceversa.

Los métodos de detección de valores atípicos (outliers) se pueden dividir en univariados y multivariados. En el caso de métodos univariados diferentes autores como Beckman & Cook (1983) tratan temas como las técnicas de rechazo para múltiples valores atípicos, valores atípicos en los datos circulares, el diseño experimental, la distribución no normal, y las series de tiempo. Barnett & Lewis (1994) hacen una unificación de los métodos de análisis estadístico para la detección de valores atípicos, así como los datos espaciales y los valores atípicos en las series temporales.

Según Gnanadesikan & Kettenring (1972), los datos atípicos multivariantes son observaciones que se consideran extrañas no por el valor que toman en una determinada variable, sino en el conjunto de aquellas. Son más difíciles de identificar que los datos atípicos unidimensionales,

dado que no pueden considerarse “valores extremos”, como sucede cuando se tiene una única variable bajo estudio.

Uno de los métodos empleados para la detección de datos atípicos multivariantes es el método de Mahalanobis o también llamado La distancia de Mahalanobis. Quaglino & Morello (2012) la definen como “un criterio muy conocido que depende de los parámetros estimados de la distribución multivariada”. Ésta describe la distancia entre cada punto de datos y el centro de masa. Cuando un punto se encuentra en el centro de masa, la distancia de Mahalanobis es cero y cuando un punto de datos se encuentra distante del centro de masa, la distancia es mayor a cero. Por lo tanto, los puntos de datos que se encuentran lejos del centro de masa se consideran valores atípicos (Matsumoto, Kamei, Monden, & Matsumoto, 2008).

La distancia de Mahalanobis se calcula para cada observación en el conjunto de datos, dándosele a cada observación un peso como inverso de la distancia de Mahalanobis. Las observaciones con valores extremos obtienen menores pesos. Finalmente, se ejecuta una regresión ponderada para minimizar el efecto de los valores extremos (Tiwari, Mehta, Jain, Tiwari, & Kanda, 2007).

Una medida de este tipo es definida de la siguiente manera:

Ecuación N.5. *Ecuación de la distancia de Mahalanobis*

$$MSD_i = \sqrt{(x_i - \bar{x})^T - S_n^{-1}(x_i - \bar{x})}$$

Dónde: x_i es el valor de la variable independiente, el súper índice T denota la matriz transpuesta, \bar{x} expresa la media del vector muestral y S_n la matriz de covarianza muestral, donde:

Ecuación N.6. *Ecuación de la matriz de covarianza muestral*

$$S_n = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)(x_i - \bar{x}_n)^T$$

Para datos multivariantes distribuidos normalmente, los valores de la distancia de Mahalanobis tienen aproximadamente una distribución chi-cuadrado con p grados de libertad. En consecuencia, aquellas observaciones con una distancia de Mahalanobis grande se indican como valores atípicos.

Los efectos de enmascaramiento y empantanamiento juegan un rol importante en la adecuación de la distancia de Mahalanobis como criterio para la detección de valores atípicos. Es decir, los efectos de enmascaramiento podrían disminuir la distancia de Mahalanobis de un valor atípico. Esto puede ocurrir, por ejemplo, cuando un pequeño grupo de datos atípicos atrae e infla hacia su dirección. Por otra parte, los efectos de empantanamiento podrían aumentar la distancia de Mahalanobis de las observaciones que no son outliers. Por ejemplo, cuando un pequeño grupo de valores atípicos atrae e infla lejos del patrón de la mayoría de las observaciones (Quaglino & Morello, 2012).

3.2.2 Estadística descriptiva

Según Rendón Macías, Villasís Kever, & Miranda Novales (2016), en su artículo “Estadística descriptiva”, la estadística descriptiva es una rama de la estadística que formula de forma clara y sencilla las recomendaciones de cómo resumir los datos de una investigación en información relevante y concisa a través de Tablas, figuras, gráficos o cuadros. El objetivo de las tablas o cuadros es proporcionar información puntual de los resultados. Las gráficas

muestran las tendencias y pueden ser histogramas, representaciones (pastel, torta), cajas con bigotes, gráficos de líneas o de puntos de dispersión. Las imágenes sirven para dar ejemplos de conceptos o reforzar hechos. La selección de un cuadro, gráfico o imagen debe basarse en los objetivos del estudio.

3.2.2.1 Promedio

De acuerdo a Levine, David, Krehbiel, Timothy, & Mark L, (2012), el promedio también conocido como media aritmética es la medida de tendencia central que más se utiliza y de mayor relevancia y representatividad en el análisis estadístico. Representa el promedio del conjunto de datos de la muestra. Su cálculo se hace sumando todos los valores de los datos, y se divide entre el número de datos que componen la muestra.

Ecuación N.7. *Ecuación matemática del promedio*

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

3.2.2.2 Mediana

Gorgas García, Cardiel López, & Zamorano Calvo (2011), definen la Mediana como una medida de tendencia central, de tal manera que, con los datos ordenados de menor a mayor, el 50% de los datos son inferiores a su valor y el 50% de los datos tienen valores superiores. Es decir, la mediana divide en dos partes iguales la distribución de frecuencias o, gráficamente, divide el histograma en dos partes de áreas iguales.

3.2.2.3 Cuartil

Los cuartiles, fraccionan a un conjunto de datos en cuatro partes iguales, existen tres cuartiles y cada una de las partes representa un 25% de los datos. El primer cuartil Q1 separa al 25%,

que abarca los valores más pequeños, del 75% restante, constituido por los valores mayores. El segundo cuartil Q2 es la mediana: 50% de sus valores son menores que la mediana y el 50% mayores que la misma. El tercer cuartil Q3 separa al 25%, que abarca a los valores más grandes, del 75% restante constituido por los que son menores (Ruíz Muñoz, 2004).

3.2.2.4 Diagrama de caja y bigotes

El diagrama de caja y bigotes ofrece una representación visual de los datos, se usa para graficar estadísticos de orden y dispersión que describen un conjunto de datos. Consiste en dibujar en eje horizontal (o vertical) un segmento de línea que va del dato menor al mayor (Rango de los datos). Entre ellos se dibujan dos rectángulos adyacentes (caja) que empieza en el valor una línea que indica Q1, le sigue mediana Q2 de los datos y termina en el valor Q3. La longitud a caja (Q3–Q1) se llama Rango intercuartil y es otra medida de dispersión de los datos.

Una gráfica de este tipo consiste en una caja rectangular, donde los lados más largos muestran el recorrido intercuartílico. Este rectángulo está dividido por un segmento vertical que indica donde se posiciona la mediana y por lo tanto su relación con los cuartiles primero y tercero (el segundo cuartil coincide con la mediana). Esta caja se ubica a escala sobre un segmento que tiene como extremos los valores mínimo y máximo de la variable. Las líneas que sobresalen de la caja se llaman bigotes. Estos bigotes tienen un límite de prolongación, de modo que cualquier dato o caso que no se encuentre dentro de este rango es marcado e identificado individualmente (Rustom Jabbaz, 2012).

3.2.3 El modelo de regresión lineal múltiple y su estimación por mínimos cuadrados.

El modelo de regresión lineal múltiple es idéntico al modelo de regresión lineal simple, con la

diferencia de que aparecen más variables explicativas:

Ecuación N.8. *Modelo de regresión lineal múltiple MCO*

$$y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + \dots + b_k \cdot x_k + u$$

En este sentido el modelo de regresión lineal múltiple es un conjunto de estimadores que logran minimizar la suma de los cuadrados de los errores entre los valores observados de la variable endógena y los que resultan del ajuste correspondiente (estimados). Este método es denotado MCO que hace referencia a las iniciales de su nombre, y consiste en hacer mínimos los errores que se producen en la estimación. Algunas de las propiedades de los estimadores son: lineales en Y, insesgados, eficientes y consistentes. (Walpole, Myers, & Myers, 1999)

3.2.4 Supuestos básicos del modelo

Para realizar un análisis de regresión lineal múltiple se hacen las siguientes consideraciones sobre los datos:

3.2.4.1 Independencia

Las perturbaciones aleatorias son independientes entre sí. Independencia entre los residuos mediante el estadístico de Durbin-Watson que toma valor 2 cuando los residuos son completamente independientes (entre 1.5 y 2.5 se considera que existe independencia), $DW < 2$ indica autocorrelación positiva y $DW > 2$ autocorrelación negativa. (Walpole et al., 1999)

Ecuación N.9. *Ecuación del supuesto de independencia entre los residuos*

$$DW = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}, \quad 0 \leq DW \leq 4$$

3.2.4.2 Normalidad de los residuos

La distribución de la perturbación aleatoria tiene distribución normal. Se puede contrastar la normalidad de los residuos mediante (Rojo Abuín, 2007):

- La prueba de Kolmogorff-Smirnov, con gráficos de normalidad de tipo Q-Q (Cuantil-Cuantil) o P-P (Proporción-Proporción).
- Gráficamente:
 - ❖ Histograma: añade una curva $N(0,1)$
 - ❖ Gráfico de probabilidad Normal de tipo P-P: Representa las proporciones acumuladas de la variable esperada respecto a las proporciones acumuladas de la variable observada.

3.2.4.3 Igualdad de varianza de los residuos (homocedasticidad)

La homocedasticidad o igualdad de varianzas de los residuos y los pronósticos. Esta condición se estudia utilizando las variables: pronósticos tipificados y Residuos tipificados mediante (Rojo Abuín, 2007):

- El estadístico de Levene.
- Gráfico de dispersión. El supuesto de homocedasticidad implica que la variación de los residuos sea uniforme en todo el rango de valores de los pronósticos (gráfico sin pautas de asociación).

3.2.4.4 Linealidad

Si no hay linealidad se dice que tenemos un error de especificación. En el caso de que sean varias variables independientes, se puede analizar esto por medio de los diagramas de dispersión parcial para cada variable independiente. En ellos se ha eliminado el efecto proveniente de las otras variables y así la relación que se muestra es la relación neta entre las variables representadas (Rojo Abuín, 2007).

3.2.4.5 Multicolinealidad o no-colinealidad

Se dice que es No-Colineal, cuando no existe colinealidad. Esta puede ser de dos clases (Walpole et al., 1999):

- Colinealidad perfecta: si una de las variables independientes tiene una relación lineal con otras independientes.
- Colinealidad parcial: si entre las variables independientes existen altas correlaciones.

4. Marco Metodológico

La estrategia metodológica para el desarrollo del proyecto está compuesta por las etapas descritas a continuación en donde se aborda de manera similar la problemática planteada en el trabajo de grado presentado por Aramburo Gonzalez en el año 2012 llamado “Diseño de herramienta para la proyección y control de la resistencia a la compresión del cemento Concretero con aplicación en una empresa cementera”. Otro insumo importante para la formulación de esta metodología fue el trabajo desarrollado por Valle García en el año 2018 llamado “Modelo de regresión lineal multivariado aplicado a la relación entre variables químicas del Clinker y variables físicas del cemento”. El trabajo se desarrolló de los siguientes pasos:

4.1 FASE I. Implementación de las ecuaciones de Bogue en los datos químicos.

En el laboratorio de calidad, se llevan muestras cada dos horas del cemento en producción, al finalizar las 24 horas de producción se juntan las muestras y se realizan los análisis químicos, físicos y mecánicos. Se toma la base de datos de 149 muestras a las que se les realiza los análisis ya mencionados y producidos a lo largo del año 2018 (desde el mes de Enero hasta Mayo) en la planta de cemento. A los datos químicos se les implementan las ecuaciones de Bogue descritas en el marco teórico para utilizarlas como variables químicas dentro de los modelos de regresión multivariada.

4.2 FASE II. Caracterización del proceso y análisis de variables.

En esta etapa se pretende realizar un análisis de estadística descriptiva para reconocer la relación entre la resistencia a 28 días y las variables físicas y químicas que se pueden estudiar en el proceso de producción. Para esto se hará uso de las herramientas descriptivas como son: la media, la mediana, la desviación estándar, el coeficiente de variación, histograma, diagrama de caja y bigotes, y la gráfica q-q para visualizar la normalidad de las variables.

4.3 FASE III. Detección de casos atípicos.

Al banco de datos se le realizará una serie de técnicas para detección de datos atípicos como por ejemplo la distancia de Mahalanobis. Con este insumo se procederá a depurar los datos para así tener una mayor confiabilidad en el modelo a desarrollar.

4.4 FASE IV. Selección de variables.

En este apartado se obtendrá la matriz de covarianza y correlación de las variables que influyen el proceso de la molienda de cemento. Estas matrices se utilizarán como insumo para identificar las variables más representativas que impactan ya sea de manera positiva o negativa la variable de interés que es la resistencia a 28 días del cemento Concretero.

4.5 FASE V. Creación del modelo de regresión.

Con la identificación de las variables más influyentes se utilizará el software Rstudio y la función “lm” para obtener el modelo regresión que describa la resistencia a 28 días del cemento Concretero en función de las variables químicas y físicas.

4.6 FASE VI. Validación del modelo.

Una vez generado el modelo se procede a validar el modelo mediante los supuestos descritos en el marco teórico en que se basa: la normalidad de los errores, la homocedasticidad o igualdad de varianza para los errores y la independencia de los errores.

5. Análisis y Resultados

En este apartado se abordan los resultados de manera ordenada con respecto a los objetivos específicos planteados:

5.1 Calcular la composición química del cemento Concretero por el método de Bogue. (C3S, C2S, C4AF, C3A)

Haciendo uso de las ecuaciones de Bogue aplicadas en el cemento y mencionadas en el marco teórico, se le realiza el cálculo de las fases químicas presentes en el cemento y resumidas en el siguiente cuadro, sin embargo, por ser una lista extensa se coloca el resto de los datos en el anexo 1:

Tabla N.2*Ecuaciones de Bogué aplicadas al cemento*

C3S	C2S	C3A	C4AF
63.98	7.44	5.77	9.92
59.35	11.59	5.68	10.29
61.16	9.97	5.7	10.38
58.74	12.03	5.57	10.53
60.31	11.21	5.47	10.71
59.54	11.68	4.92	10.92
61.59	8.81	4.67	11.23
60.29	9.94	5.04	11.2
61.43	9.39	5.45	10.74
64.43	6.64	5.63	10.32
65.68	5.98	5.56	10.01
65.61	5.98	5.53	9.89
63.05	8.74	5.07	10.29
66.03	5.87	5.52	10.13
64.43	7.3	5.65	10.04
65.41	6.1	5.25	10.38
63.27	7.92	5.28	10.47
66.41	5.64	5	10.25
66.19	5.57	5.03	10.07
57.16	13.45	5.16	10.65
63.47	8.37	5.29	10.22
63.46	8.43	5.45	10.41
61.89	9.53	5.63	10.89

5.2 Realizar análisis descriptivo de las variables que influyen en la resistencia a 28 días en el cemento Concretero.

De acuerdo con el proceso de molienda de cemento, se han definido las siguientes variables para analizar su influencia en la resistencia a 28 días del cemento Concretero:

Tabla N. 3*Variables que influyen en la resistencia a 28 días*

Propiedad	Nombre de la variable
% de partículas menor a 3 μ m	d3 μ m
% de partículas entre 3 μ m y 30 μ m	d3_d30 μ m
% de partículas mayor a 30 μ m	d30 μ m

Blaine (cm2/g)	Blaine
Retenido en malla 325	Ret_325
Alita	C3S
Belita	C2S
Celita	C3A
Felita	C4AF
Pérdida ígnea	P_Ignea
Cal libre	Cal_libre
Relación Sodio – Potasio	EA
Resistencia a 28 días (Mpa)	R28

Calculando la distancia de Mahalanobi para los datos, se tiene lo siguiente:

[1]	10.114791	7.015615	7.734558	6.985357	13.512119	19.881968
[7]	16.727628	16.306231	6.460494	11.296639	12.789106	15.328436
[13]	10.912026	9.147192	21.336093	10.059581	8.798760	11.582724
[19]	11.584798	10.497703	15.987878	16.216093	14.091901	11.551283
[25]	18.220484	13.550243	10.320432	6.030737	4.704120	40.096402
[31]	20.946855	9.672432	14.764985	8.631493	8.514419	11.233367
[37]	13.985947	10.651609	13.604560	6.720505	11.432627	7.665401
[43]	6.926099	11.614011	13.457148	12.050353	16.792067	11.859265
[49]	10.293951	13.611939	15.803964	16.605107	15.162389	7.673893
[55]	11.837706	9.955860	18.877865	19.089779	17.404427	16.387730
[61]	18.249770	17.702871	14.406077	11.181082	13.858106	15.468710
[67]	11.511548	10.202828	11.517853	10.896724	11.091282	4.783105
[73]	10.501175	24.947097	16.933762	10.629251	11.589363	125.621840
[79]	15.056516	13.567986	9.415191	10.985948	14.278812	13.379436
[85]	23.334357	12.537830	14.047930	11.905908	18.779485	19.705621
[91]	14.753243	15.371890	10.084786	8.455970	7.811296	7.871339
[97]	8.467150	14.785717	25.988070	10.450549	7.723659	6.206199
[103]	9.650008	19.623355	10.474272	13.087646	12.789952	9.650105
[109]	5.355471	23.393479	14.152870	11.332947	15.309186	9.428180
[115]	12.577027	10.514406	5.760058	12.930998	11.047329	10.071487
[121]	9.402171	7.534707	6.835213	13.417914	4.663972	10.414471
[127]	16.822923	18.622314	12.372706	4.110224	11.488786	9.679789
[133]	8.310438	6.258949	4.288503	3.080510	5.288764	10.444274
[139]	10.040642	7.516814	5.206507	14.138877	10.658471	10.372283
[145]	9.680391	7.517012	6.846850	13.690155	10.020116	

Figura N.1. Datos en RStudio de la distancia de Mahalanobi

Según este resultado las muestras con una distancia mayor a 20, son valores atípicos, a saber: 15, 30, 31, 74, 78, 85, 99 y 110. Por ello se sacan de la base de datos y se trabaja con las 141 muestras restantes.

Utilizando el programa Rstudio, se obtienen las siguientes tablas con los siguientes estadísticos:

Tabla N. 4

Estadísticos de las primeras 5 variables

	d3 μ m	d3_30 μ m	d30 μ m	Blaine	Ret_325	R28
Valor mínimo	14.98	67.42	0.00	4168	0.16	43.10
1er Cuartil	17.99	72.99	2.23	4599	0.32	46.70
Mediana	19.23	74.10	6.82	4710	0.38	48.70
Promedio	19.71	74.69	5.61	4753	0.38	48.60
3er Cuartil	21.25	76.20	8.64	4903	0.45	50.50
Valor máximo	27.38	81.27	14.86	5675	0.62	54.50
CV	0.12	0.03	0.66	0.06	0.22	0.05
Asimetría	0.82	0.44	-0.12	0.72	0.13	0.06
Curtosis	3.51	3.18	1.85	4.10	2.74	2.37
Valor P (Normalidad Shapiro - Wilk test)	0.000	0.001	0.000	0.001	0.868	0.322

Tabla N. 5

Estadísticos de las siguientes 5 variables

	C3S	C2S	C3A	C4AF	P_lgnea	Cal_libre	EA
Valor mínimo	52.19	0.11	4.67	8.64	2.66	0.58	0.15
1er Cuartil	59.91	6.60	5.53	9.31	3.90	0.92	0.29
Mediana	62.28	8.54	6.02	9.68	4.67	1.13	0.33
Promedio	62.10	8.68	5.91	9.74	4.52	1.17	0.33
3er Cuartil	64.11	11.03	6.26	10.22	5.16	1.37	0.37
Valor máximo	72.32	18.05	6.77	11.23	6.17	2.18	0.55
CV	0.06	0.38	0.08	0.06	0.18	0.29	0.21
Asimetría	-0.05	0.20	-0.33	0.35	-0.25	0.69	0.43
Curtosis	3.29	3.20	2.35	2.49	2.15	3.21	4.01
Valor P (Normalidad Shapiro - Wilk test)	0.764	0.527	0.007	0.022	0.005	0.001	0.010

A continuación, se realiza un gráfico de caja y bigotes, un histograma y un gráfico Q- Q para visualizar el comportamiento de cada una de las variables.

- *% de partículas menor a $3\mu\text{m}$*

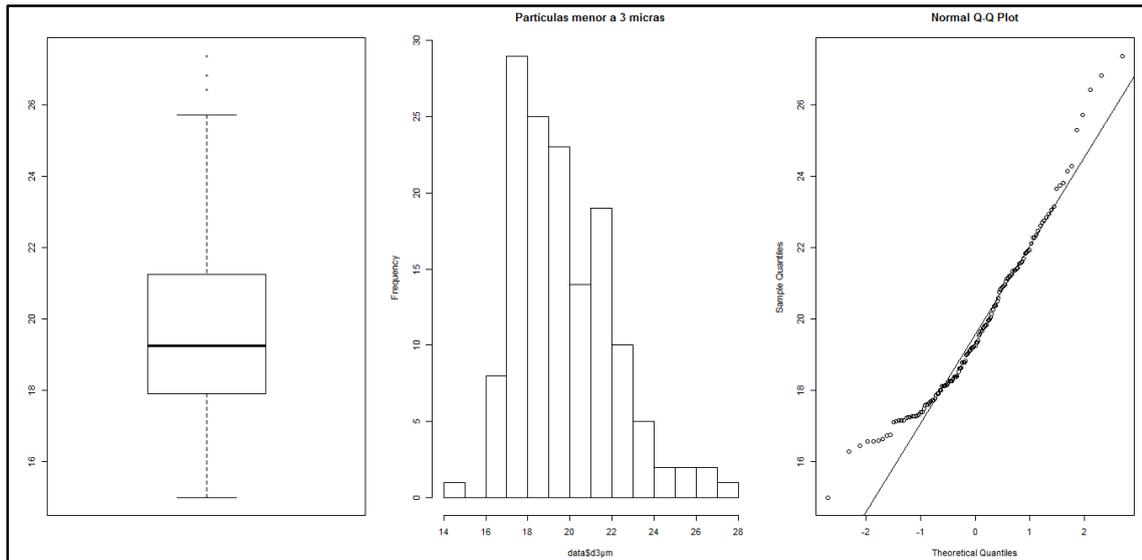


Figura N.2. Diagrama en RStudio % de partículas menor a $3\mu\text{m}$

Como lo muestra la tabla 4 y la figura 2, esta variable tiene una variación baja de tan sólo 12%, presenta asimetría hacia la derecha, tiene una forma leptocúrtica, se puede decir que no tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes los siguientes valores se tienen por fuera: 27.38; 26.83; 26.42

- *% de partículas entre $3\mu\text{m}$ y $30\mu\text{m}$*

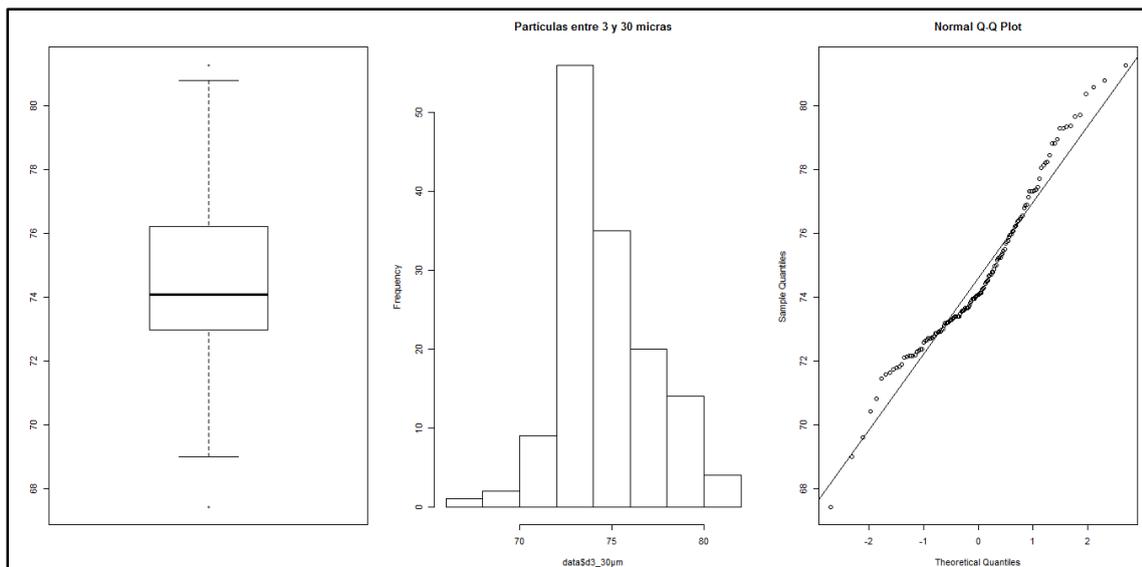


Figura N.3. Diagrama en RStudio % de partículas entre $3\mu\text{m}$ y $30\mu\text{m}$

Como lo muestra la tabla 4 y la figura 3, esta variable tiene una variación muy baja de tan

sólo 3%, presenta asimetría hacia la derecha, tiene una forma leptocúrtica, se puede decir que no tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes los siguientes valores se tienen por fuera: 81.27; 67.42

- *% de partículas mayor a 30 μ m*

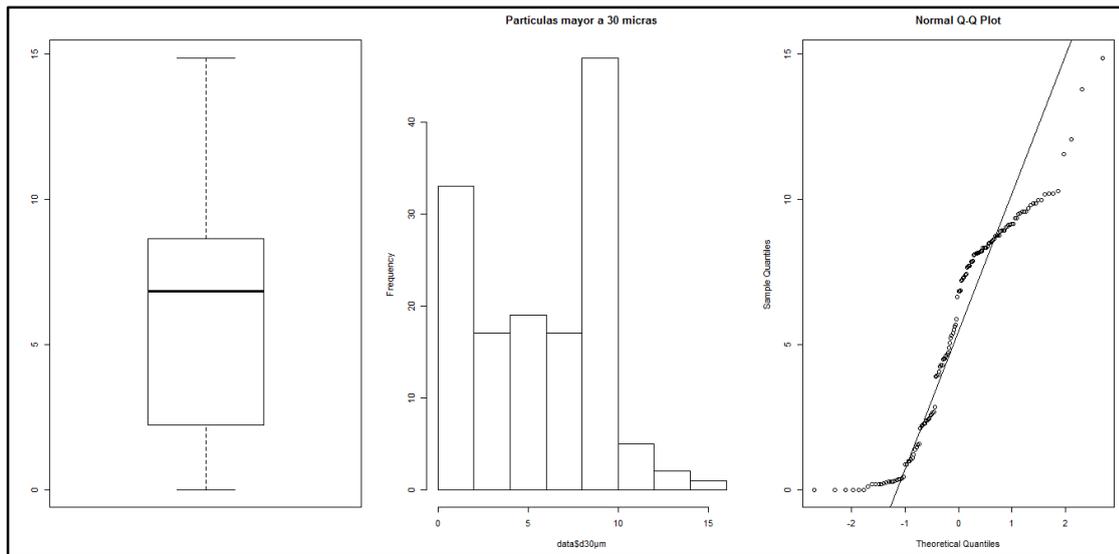


Figura N.4. Diagrama en RStudio % de partículas mayor 30 μ m

Como lo muestra la tabla 4 y la figura 4, esta variable tiene una variación muy alta del 66%, presenta asimetría hacia la izquierda, tiene una forma leptocúrtica, se puede decir que no tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes no registra valores atípicos.

- *Blaine (cm²/g)*

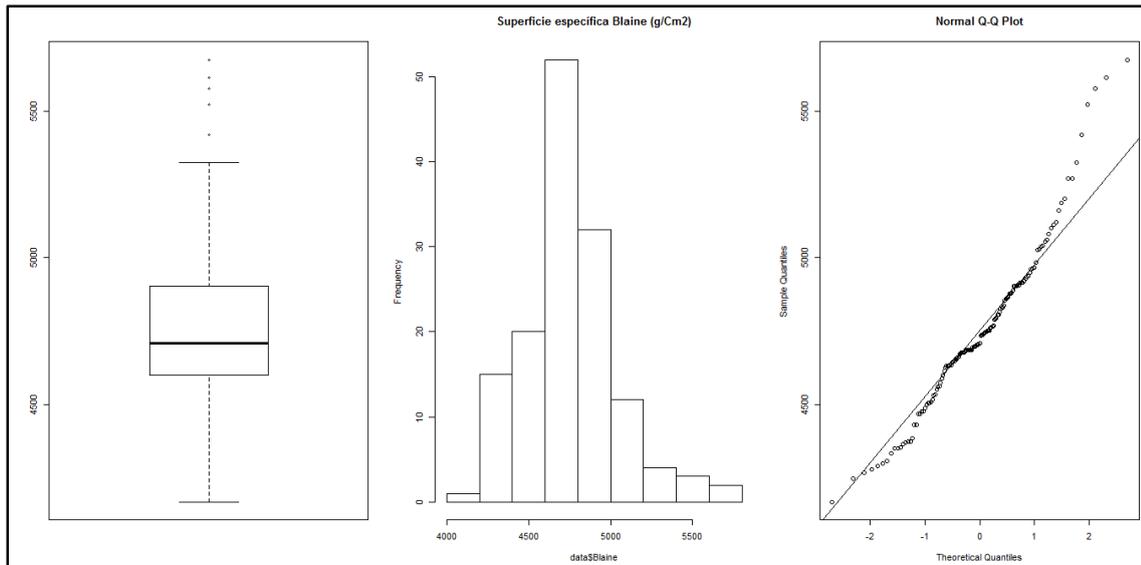


Figura N.5. Diagrama en RStudio Blaine (cm²/g)

Como lo muestra la tabla 4 y la figura 5, esta variable tiene una variación muy baja de tan sólo 6%, presenta asimetría hacia la derecha, tiene una forma leptocúrtica, se puede decir que no tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk.

En el diagrama de caja y bigotes los siguientes valores se tienen por fuera (datos atípicos):

5675; 5418; 5521; 5614; 5575.

- *Retenido en malla 325*

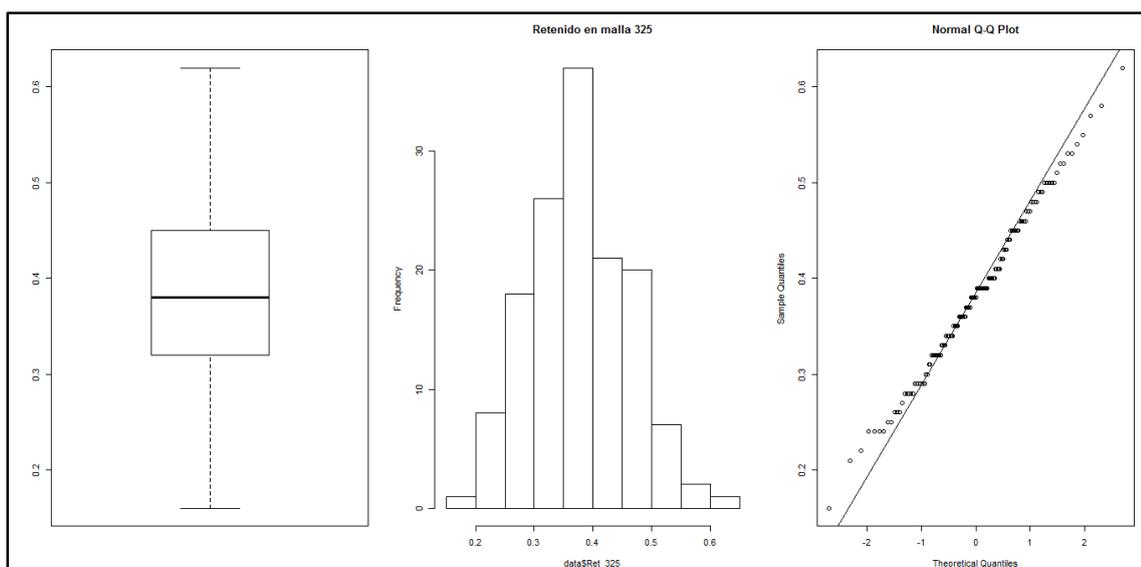


Figura N.6. Diagrama en RStudio retenido en malla 325

Como lo muestra la tabla 4 y la figura 6, esta variable tiene una variación baja con respecto al Blaine (22%), presenta asimetría hacia la derecha, tiene una forma leptocúrtica, se puede decir que tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes no registra valores atípicos.

- *Resistencia a 28 días*

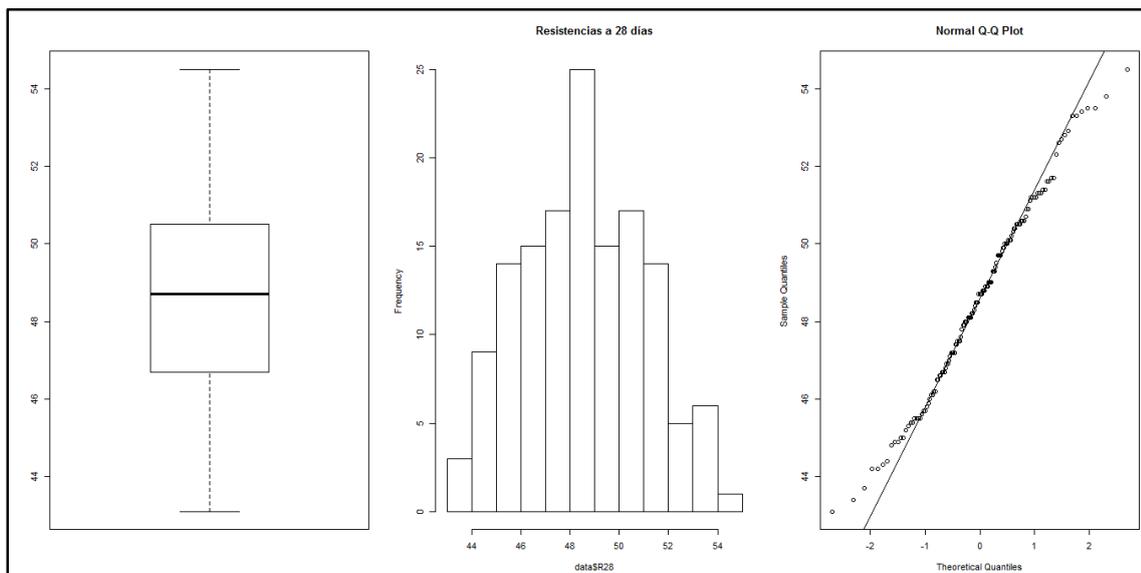


Figura N.7. Diagrama en RStudio Resistencia a 28 días

Como lo muestra la tabla 4 y la figura 7, esta variable tiene una variación baja de tan sólo 5%, esto es bueno puesto que será la variable regresora. Presenta una leve asimetría hacia la derecha, tiene una forma leptocúrtica, se puede decir que tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes no registra valores atípicos.

- *Contenido de Alita (C₃S)*

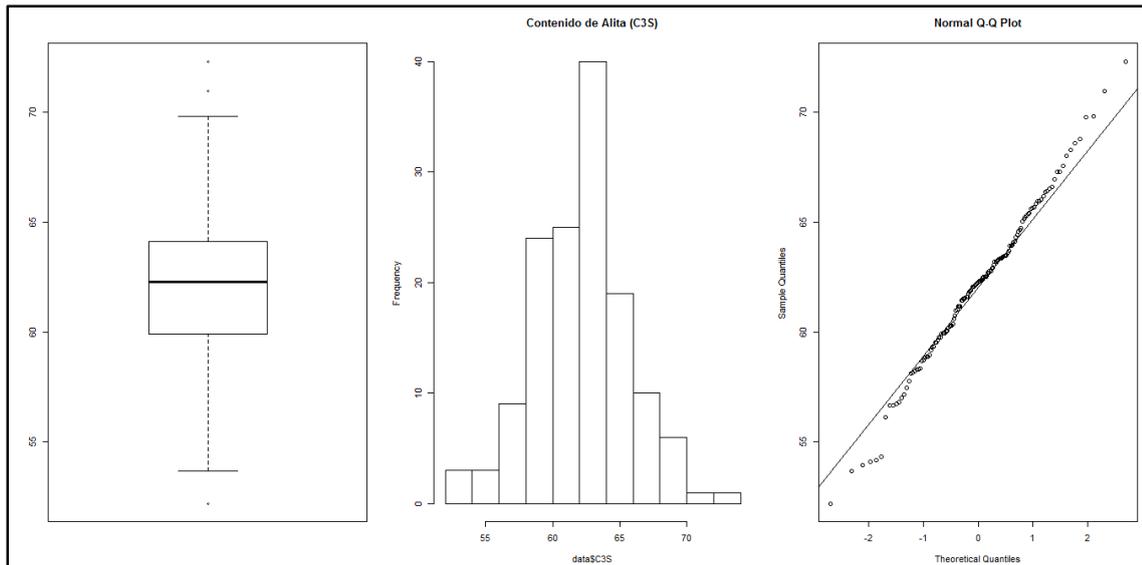


Figura N.8. Diagrama en RStudio Contenido de Alita C3S

Como lo muestra la tabla 5 y la figura 8, esta variable tiene una variación muy baja de tan sólo 6%, presenta asimetría hacia la izquierda, tiene una forma leptocúrtica, se puede decir que tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes los siguientes valores se tienen por fuera (datos atípicos): 70.96; 72.32; 52.19.

- *Contenido de Belita (C₂S)*

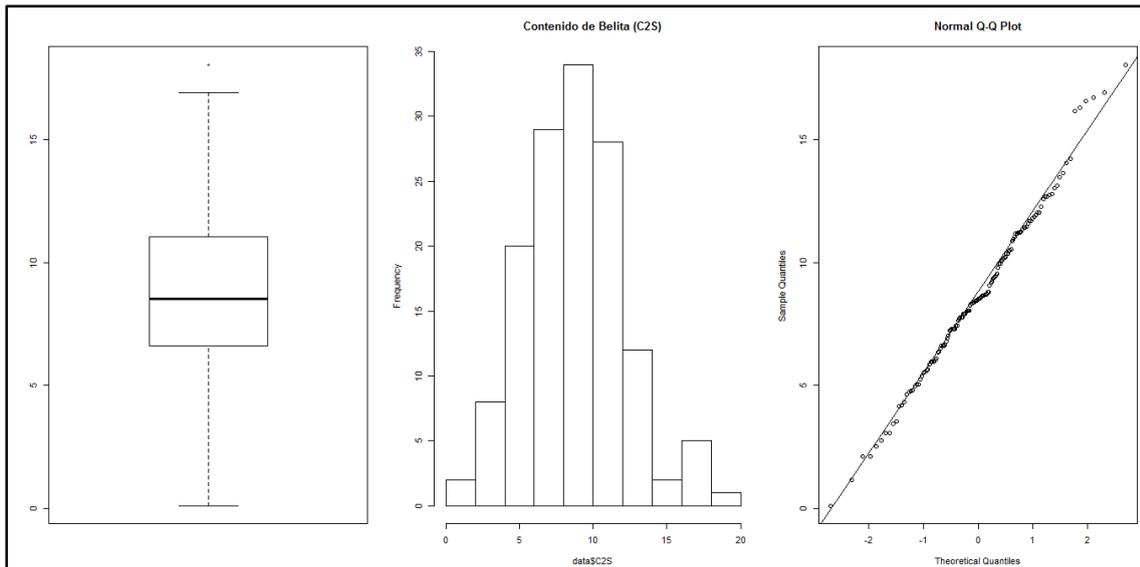


Figura N.9. Diagrama en RStudio contenido de Belita C2S

Como lo muestra la tabla 5 y la figura 9, esta variable tiene una variación en comparación con las otras variables un poco alta de 38%, presenta asimetría hacia la derecha, tiene una forma leptocúrtica, se puede decir que tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes sólo presenta un valor atípico: 18.05.

- *Contenido de Celita(C3A)*

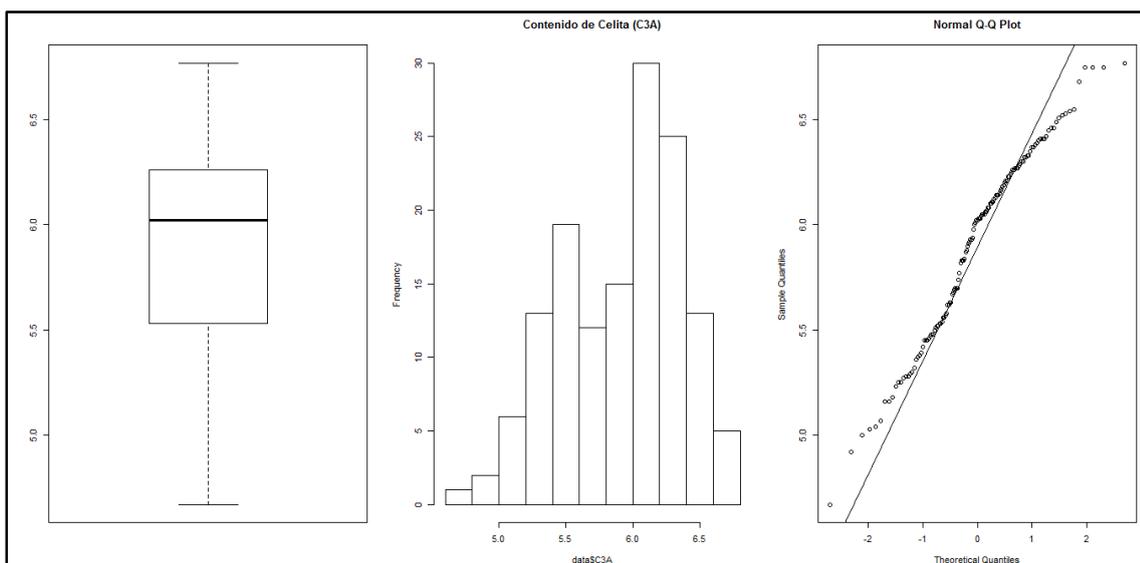


Figura N.10. Diagrama en RStudio contenido de Celita C3A

Como lo muestra la tabla 5 y la figura 10, esta variable también tiene una variación muy baja

de tan sólo 8%, presenta asimetría hacia la izquierda, tiene una forma leptocúrtica, se puede decir que no tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes no registra valores atípicos.

- *Contenido de Felita (C4AF)*

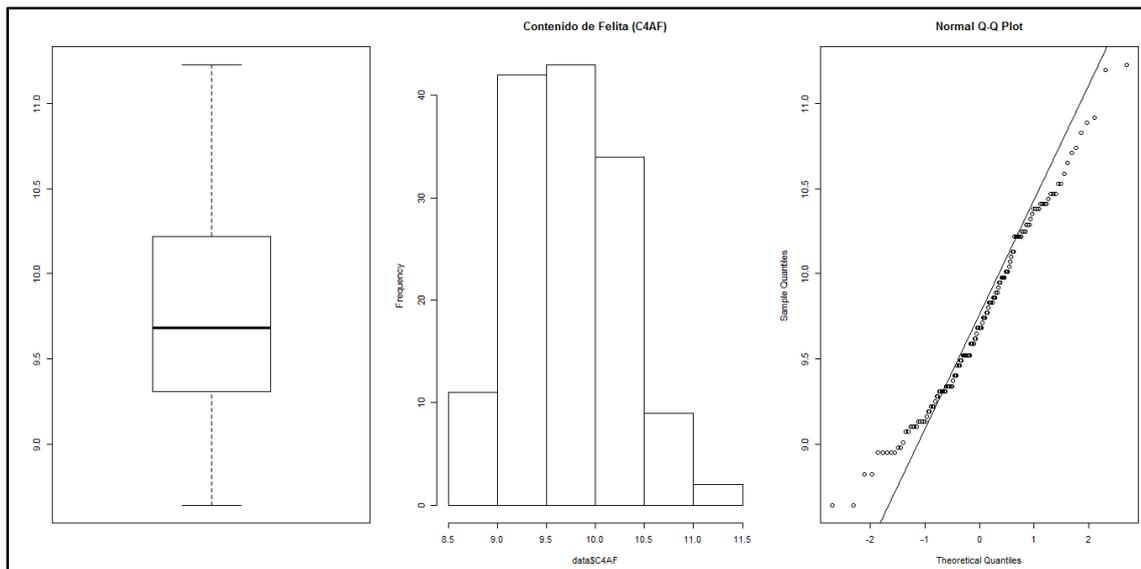


Figura N.11. Diagrama en RStudio contenido de Felita C4AF

Como lo muestra la tabla 5 y la figura 11, esta variable tiene una variación muy baja igual al del C₃S de tan sólo 6%, presenta asimetría hacia la derecha, tiene una forma leptocúrtica, se puede decir que no tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes no registra valores atípicos.

- *Pérdida ígnea (P_Ignea)*

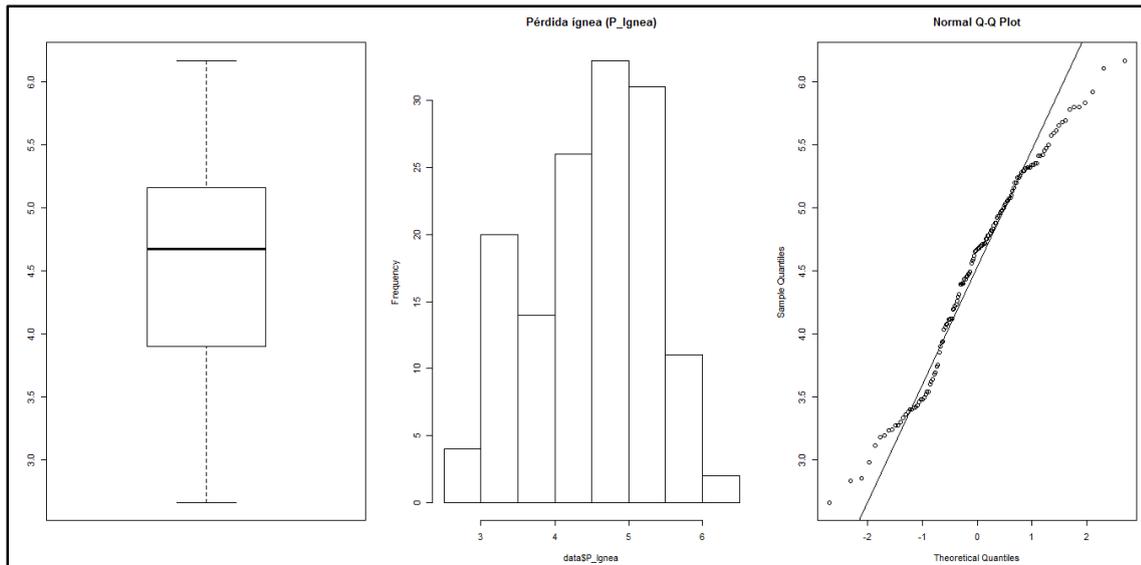


Figura N.12. Diagrama en RStudio pérdida ígnea

Como lo muestra la tabla 5 y la figura 12, esta variable tiene una variación relativamente baja pero de tan sólo 18%, presenta asimetría hacia la izquierda, tiene una forma leptocúrtica, se puede decir que no tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes no registra valores atípicos.

- *Cal libre (Cal_libre)*

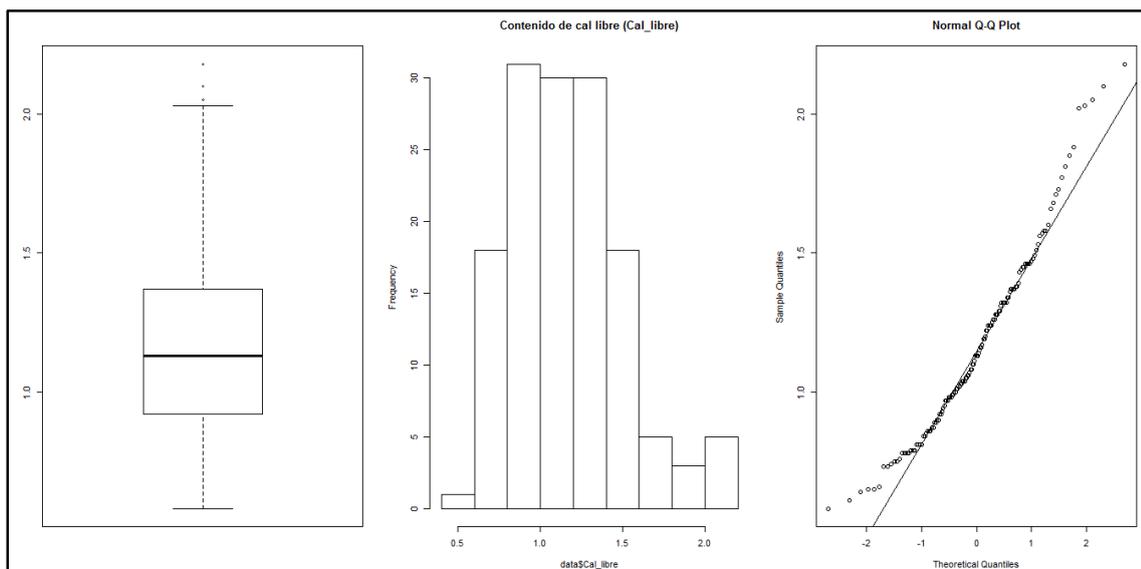


Figura N.13. Diagrama en RStudio de Cal libre

Como lo muestra la tabla 5 y la figura 9, esta variable tiene una variación medianamente baja

del 29%, presenta asimetría hacia la derecha, tiene una forma leptocúrtica, se puede decir que no tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes registra los siguientes valores atípicos: 2.18; 2.05; 2.10.

- *Relación Sodio – Potasio (EA)*

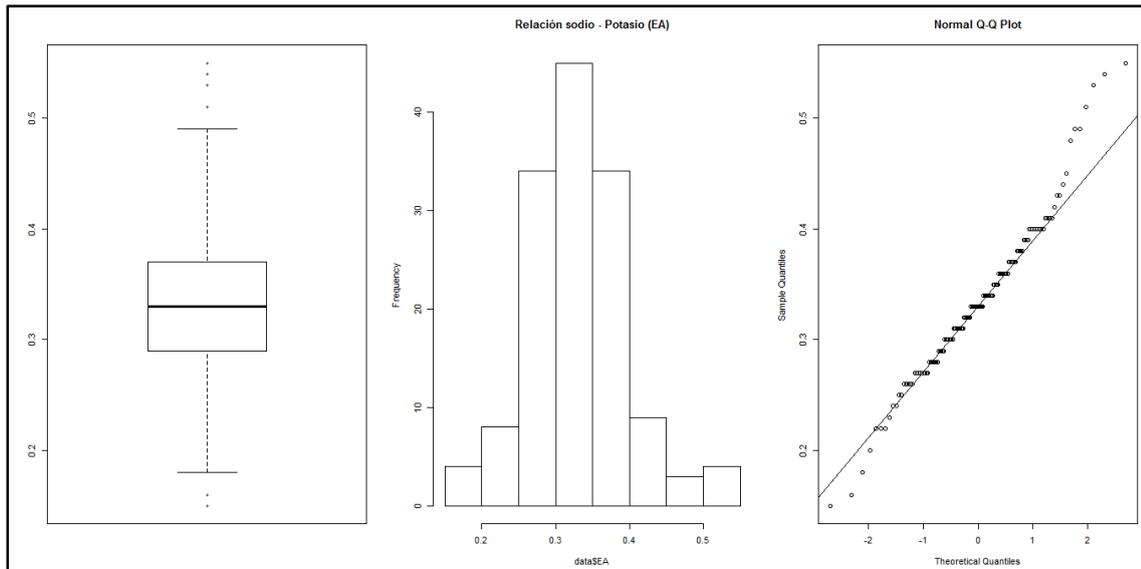


Figura N.14. Diagrama en RStudio relación Sodio-Potasio EA

Como lo muestra la tabla 5 y la figura 14, esta variable tiene una variación medianamente baja y similar a la de la pérdida ígnea del 21%, presenta asimetría hacia la derecha, tiene una forma leptocúrtica, se puede decir que no tiene una distribución normal tal como lo muestra el gráfico Q-Q y el test de Shapiro-Wilk. En el diagrama de caja y bigotes registra los siguientes valores atípicos: 0.15; 0.54; 0.16; 0.51; 0.53; 0.55.

5.3 Realizar análisis multivariado para elegir las variables de mayor

influencia en la resistencia a 28 días.

- Matriz de Covarianza

	d3µm	d3_30µm	d30µm	Blaine	Ret_325
d3µm	5.56951317	0.91396594	-6.48347911	232.0434372	-0.0103428116
d3_30µm	0.91396594	6.24164767	-7.15561361	-44.9155446	-0.0253894377
d30µm	-6.48347911	-7.15561361	13.63909272	-187.1278926	0.0357322492
Blaine	232.04343718	-44.91554458	-187.12789260	77975.7411348	-6.7305091185
Ret_325	-0.01034281	-0.02538944	0.03573225	-6.7305091	0.0072561702
R28	-1.03165714	-0.94405000	1.97570714	-299.1485714	-0.0474928571
C3S	1.78775789	0.07527728	-1.86303517	-69.7828632	-0.0213968389
C2S	-1.70519051	-0.06372861	1.76891912	-14.3265405	0.0196372644
C3A	-0.22226056	0.08396511	0.13829545	45.7924686	-0.0076467629
C4AF	-0.06488888	-0.32392937	0.38881825	-81.9751525	0.0077695289
P_Ignea	0.27967540	0.12134763	-0.40102303	161.3243485	-0.0077450456
Cal_libre	0.23530932	-0.12157058	-0.11373874	-0.3617204	0.0009624468
EA	0.02037092	0.05788010	-0.07825102	2.3640299	-0.0002953040

Figura N.15. Datos en RStudio Matriz de covarianza d3µm, d3-30µm, d30µm, Blaine, Ret325

	R28	C3S	C2S	C3A	C4AF
d3µm	-1.03165714	1.78775789	-1.70519051	-0.222260557	-0.064888880
d3_30µm	-0.94405000	0.07527728	-0.06372861	0.083965111	-0.323929367
d30µm	1.97570714	-1.86303517	1.76891912	0.138295446	0.388818247
Blaine	-299.14857143	-69.78286322	-14.32654053	45.792468592	-81.975152482
Ret_325	-0.04749286	-0.02139684	0.01963726	-0.007646763	0.007769529
R28	6.32442857	1.80700000	-1.36915000	-0.277485714	0.735692857
C3S	1.80700000	13.01987687	-11.86893111	-0.558031413	0.262867781
C2S	-1.36915000	-11.86893111	11.17088737	0.432017958	-0.128692411
C3A	-0.27748571	-0.55803141	0.43201796	0.203280436	-0.186184797
C4AF	0.73569286	0.26286778	-0.12869241	-0.186184797	0.305743749
P_Ignea	-1.01460714	-0.23357789	-0.13416877	0.188064843	-0.334724691
Cal_libre	0.16503571	0.35846758	-0.29436754	-0.075537128	0.068920638
EA	-0.05317857	0.02753384	-0.03518770	0.001544032	-0.009231829

Figura N.16. Datos en RStudio Matriz de covarianza R28, C3S, C2S, C3A, C4AF

	P_Ignea	Cal_libre	EA
d3µm	0.279675400	0.2353093161	0.020370922
d3_30µm	0.121347634	-0.1215705775	0.057880101
d30µm	-0.401023034	-0.1137387386	-0.078251023
Blaine	161.324348531	-0.3617203647	2.364029889
Ret_325	-0.007745046	0.0009624468	-0.000295304
R28	-1.014607143	0.1650357143	-0.053178571
C3S	-0.233577888	0.3584675836	0.027533845
C2S	-0.134168774	-0.2943675380	-0.035187700
C3A	0.188064843	-0.0755371277	0.001544032
C4AF	-0.334724691	0.0689206383	-0.009231829
P_Ignea	0.660310314	-0.0906271733	0.015752649
Cal_libre	-0.090627173	0.1143354407	-0.002807584
EA	0.015752649	-0.0028075836	0.004853536

Figura N.17. Datos en RStudio Matriz de covarianza Pérdida ígnea, Cal libre, EA

- Matriz de correlación

	d3 μ m	d3_30 μ m	d30 μ m	Blaine	Ret_325	R28
d3 μ m	1.00000000	0.155014301	-0.74388664	0.352112073	-0.05144892	-0.1738266
d3_30 μ m	0.15501430	1.00000000	-0.77554056	-0.064382419	-0.11930255	-0.1502570
d30 μ m	-0.74388664	-0.775540559	1.00000000	-0.181453784	0.11358314	0.2127253
Blaine	0.35211207	-0.064382419	-0.18145378	1.00000000	-0.28295308	-0.4259872
Ret_325	-0.05144892	-0.119302550	0.11358314	-0.282953083	1.00000000	-0.2216991
R28	-0.17382656	-0.150257004	0.21272532	-0.425987182	-0.22169910	1.0000000
C3S	0.20994070	0.008350469	-0.13980570	-0.069257305	-0.06961334	0.1991334
C2S	-0.21618246	-0.007632046	0.14330832	-0.015350331	0.06897368	-0.1628908
C3A	-0.20888439	0.074542071	0.08305530	0.363719436	-0.19910222	-0.2447272
C4AF	-0.04972592	-0.234488692	0.19040358	-0.530913397	0.16495374	0.5290626
P_Ignea	0.14583831	0.059773357	-0.13362956	0.710961566	-0.11189127	-0.4964931
Cal_libre	0.29487644	-0.143909154	-0.09108046	-0.003830917	0.03341432	0.1940782
EA	0.12390047	0.332545084	-0.30413623	0.121518970	-0.04976069	-0.3035267

Figura N.18. Datos en RStudio Matriz de correlación d3 μ m, d3-30 μ m, d30 μ m, Blaine, Ret325, R28

	C3S	C2S	C3A	C4AF	P_Ignea	Cal_libre
d3 μ m	0.209940703	-0.216182459	-0.20888439	-0.04972592	0.14583831	0.294876436
d3_30 μ m	0.008350469	-0.007632046	0.07454207	-0.23448869	0.05977336	-0.143909154
d30 μ m	-0.139805698	0.143308318	0.08305530	0.19040358	-0.13362956	-0.091080463
Blaine	-0.069257305	-0.015350331	0.36371944	-0.53091340	0.71096157	-0.003830917
Ret_325	-0.069613336	0.068973684	-0.19910222	0.16495374	-0.11189127	0.033414316
R28	0.199133381	-0.162890797	-0.24472717	0.52906265	-0.49649311	0.194078245
C3S	1.000000000	-0.984156883	-0.34301050	0.13175137	-0.07966256	0.293802857
C2S	-0.984156883	1.000000000	0.28668843	-0.06963544	-0.04940075	-0.260468761
C3A	-0.343010502	0.286688434	1.00000000	-0.74682291	0.51331716	-0.495475563
C4AF	0.131751366	-0.069635438	-0.74682291	1.00000000	-0.74496336	0.368620991
P_Ignea	-0.079662558	-0.049400754	0.51331716	-0.74496336	1.00000000	-0.329832889
Cal_libre	0.293802857	-0.260468761	-0.49547556	0.36862099	-0.32983289	1.000000000
EA	0.109530305	-0.151118599	0.04915634	-0.23965132	0.27825978	-0.119182676

Figura N.19. Datos en RStudio Matriz de correlación C3S, C2S, C3A, C4AF, Pérdida ígnea, cal libre.

	EA
d3 μ m	0.12390047
d3_30 μ m	0.33254508
d30 μ m	-0.30413623
Blaine	0.12151897
Ret_325	-0.04976069
R28	-0.30352670
C3S	0.10953031
C2S	-0.15111860
C3A	0.04915634
C4AF	-0.23965132
P_Ignea	0.27825978
Cal_libre	-0.11918268
EA	1.00000000

Figura N.20. Datos en RStudio Matriz de correlación EA

- Diagramas de dispersión

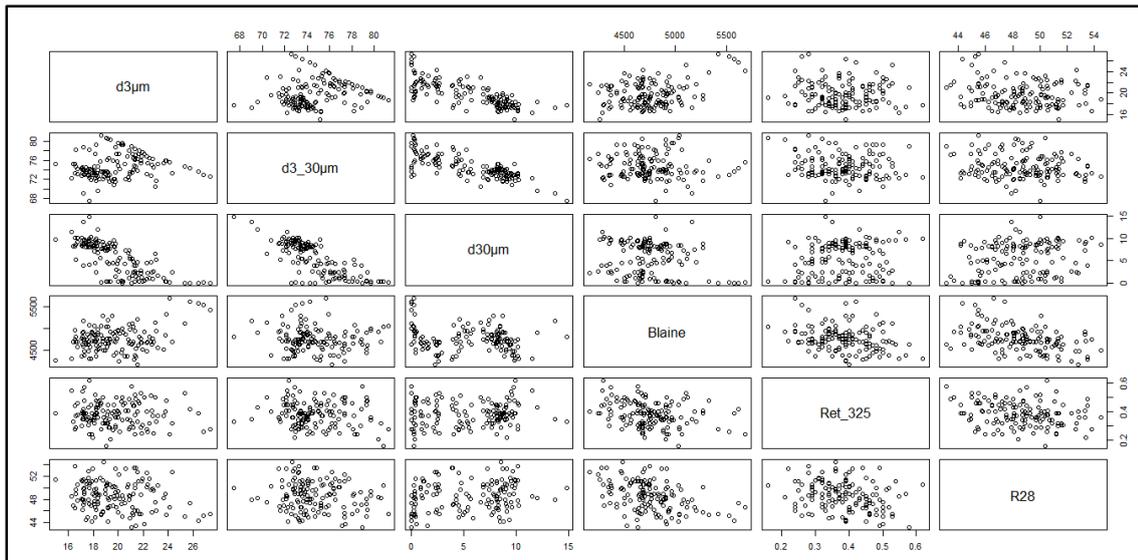


Figura N.21. Diagrama de dispersión $d_{3\mu m}$, $d_{3-30\mu m}$, $d_{30\mu m}$, Blaine, Ret₃₂₅, R₂₈

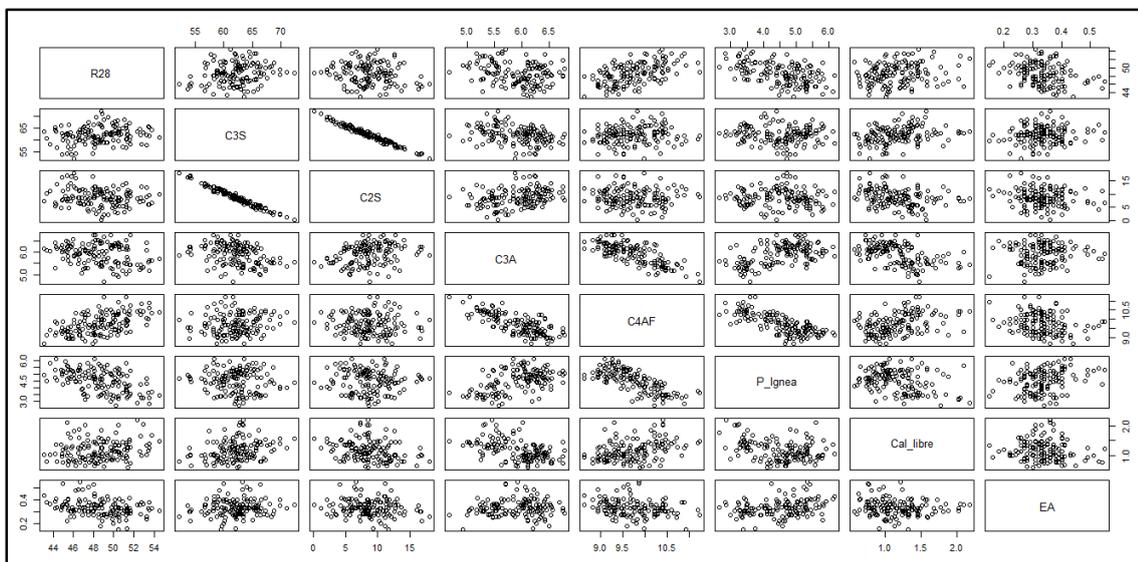


Figura N.22. Diagrama de dispersión R₂₈, C_{2S}, C_{3S}, C_{3A}, C_{4AF}, Pérdida ígnea, cal libre, EA

Como lo muestran las figura 20 y 22, es de esperarse una correlación alta entre el C_{3S} y C_{2S} por su relación matemática lineal.

- *Selección de variables:* Se evalúan dos criterios de selección.

1. Valor P: Para un nivel crítico de significancia del 5%, se eliminan todas las variables con valor P encima del 5%. El problema de utilizar este primer criterio es que usa un criterio de selección de variables demasiado estricto.

A menudo al usarlo se desperdicia información útil para pronosticar la variable dependiente, solo porque no se puede establecer una relación estadísticamente significativa. Pero la definición de “estadísticamente significativa” es arbitraria, puesto que no se puede decir que una variable con valor $-p$ de 0.028 es mejor que una con 0.061.

2. R-Cuadrada ajustada (coeficiente de correlación ajustado): Con este criterio se puede decidir si la variable incluida o eliminada contribuye o no al modelo de regresión.

A continuación, se ilustra el primer modelo de regresión lineal corrido en R:

```
Call:
lm(formula = R28 ~ ., data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.1006 -1.1755  0.1788  1.0678  4.2629

Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.467e+01  3.657e+01   2.588 0.010750 *
d3µm         -1.032e-02  7.730e-02  -0.134 0.893992
d3_30µm      -4.079e-02  6.983e-02  -0.584 0.560170
d30µm        NA          NA          NA    NA
Blaine       -3.090e-03  9.306e-04  -3.321 0.001167 **
Ret_325      -1.066e+01  1.923e+00  -5.544 1.6e-07 ***
C3S          -7.053e-01  3.944e-01  -1.788 0.076129 .
C2S          -9.062e-01  4.240e-01  -2.137 0.034449 *
C3A           1.622e+00  6.249e-01   2.595 0.010544 *
C4AF          2.206e+00  6.422e-01   3.436 0.000795 ***
P_Ignea      -4.267e-01  5.083e-01  -0.840 0.402736
Ca_libre      6.175e-01  5.757e-01   1.073 0.285442
EA           -6.716e+00  2.397e+00  -2.802 0.005861 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.747 on 129 degrees of freedom
Multiple R-squared:  0.5552,    Adjusted R-squared:  0.5173
F-statistic: 14.64 on 11 and 129 DF,  p-value: < 2.2e-16
```

Figura N.23. Datos en RStudio modelo de regresión lineal

Para este primer modelo se visualiza que hay una variable que no converge en la regresión

(d30 μ m) y dos variables (d3 μ m y d3_30 μ m) con valor-p muy grandes en comparación a las otras. Con un R cuadrado ajustado del modelo de 0,5173. A continuación, se vuelve a correr el modelo sin estas tres variables para revisar cómo se mueven los valores-p.

```
Call:
lm(formula = R28 ~ Blaine + Ret_325 + C3S + C2S + C3A + C4AF +
    P_Ignea + EA, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.1101 -1.3248  0.2691  0.9931  4.4322

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.438e+01  3.398e+01   2.777 0.006282 **
Blaine       -2.684e-03  7.924e-04  -3.387 0.000932 ***
Ret_325      -1.038e+01  1.880e+00  -5.520 1.73e-07 ***
C3S          -7.449e-01  3.851e-01  -1.934 0.055201 .
C2S          -9.609e-01  4.132e-01  -2.325 0.021588 *
C3A           1.485e+00  5.633e-01   2.636 0.009387 **
C4AF          2.242e+00  6.004e-01   3.734 0.000279 ***
P_Ignea      -5.719e-01  4.706e-01  -1.215 0.226372
EA           -7.374e+00  2.284e+00  -3.228 0.001572 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.738 on 132 degrees of freedom
Multiple R-squared:  0.5494,    Adjusted R-squared:  0.5221
F-statistic: 20.12 on 8 and 132 DF,  p-value: < 2.2e-16
```

Figura N.24. Datos en RStudio, segundo modelo de regresión lineal

En este segundo modelo el valor-p de la variable P_Ignea sigue siendo mayor al de la significancia definida, el valor R cuadrado ajustado subió a 0.5221. Con esto se puede decir que la variable eliminada no contribuía al modelo. Sin embargo, a continuación, se corre nuevamente el modelo sin la variable P_Ignea para revisar el impacto que tiene:

```

Call:
lm(formula = R28 ~ Blaine + Ret_325 + C3S + C2S + C3A + C4AF +
    EA, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.2104 -1.2999  0.1297  1.0224  4.3599

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.579e+01  2.458e+01   2.677  0.00836 **
Blaine       -3.118e-03  7.083e-04  -4.402  2.18e-05 ***
Ret_325     -1.053e+01  1.879e+00  -5.601  1.17e-07 ***
C3S         -4.231e-01  2.801e-01  -1.510  0.13331
C2S         -6.124e-01  2.981e-01  -2.054  0.04192 *
C3A          1.608e+00  5.551e-01  2.897  0.00441 **
C4AF         2.696e+00  4.705e-01  5.731  6.37e-08 ***
EA          -7.500e+00  2.286e+00  -3.281  0.00132 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.742 on 133 degrees of freedom
Multiple R-squared:  0.5444,    Adjusted R-squared:  0.5204
F-statistic: 22.7 on 7 and 133 DF,  p-value: < 2.2e-16

```

Figura N.25. Datos en RStudio segundo modelo de regresión lineal sin la variable Pérdida ígnea

Como se puede ver cuando se elimina la variable P_Ígnea el modelo pierde bondad de ajuste como lo muestra el R cuadrado ajustado que pasa de 0.5221 a 0.5204. Sin embargo no es considerable el aporte al modelo.

Por otro lado, es necesario eliminar una de las variables de composición química ya sea el C₃S o C₂S puesto que son definidas por una relación matemática una en función de la otra y esto a su vez puede traer problemas de colinealidad al modelo dado que el coeficiente de correlación es de 0.98. Se elimina del modelo la variable C₃S por tener un valor-p superior al del C₂S, arrojando el siguiente modelo:

```

Call:
lm(formula = R28 ~ Blaine + Ret_325 + C2S + C3A + C4AF + EA,
    data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.1158 -1.3306  0.1399  1.2266  4.2931

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 31.1903526   8.9375361   3.490 0.000655 ***
Blaine      -0.0026683   0.0006456  -4.133 6.28e-05 ***
Ret_325     -9.9010630   1.8418932  -5.375 3.29e-07 ***
C2S         -0.1679349   0.0481290  -3.489 0.000656 ***
C3A          1.8058159   0.5419613   3.332 0.001115 **
C4AF         2.7664126   0.4704149   5.881 3.09e-08 ***
EA          -6.7894838   2.2476216  -3.021 0.003021 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.75 on 134 degrees of freedom
Multiple R-squared:  0.5366,    Adjusted R-squared:  0.5158
F-statistic: 25.86 on 6 and 134 DF,  p-value: < 2.2e-16

```

Figura N.26. Datos en RStudio Modelo de regresión lineal sin C3S

Como se puede ver el modelo pierde algo la bondad de ajuste, sin embargo, por lo anterior es necesario eliminar la variable. A continuación, se presenta el modelo elegido:

$$R28 = \text{Intercepto} + \text{Blaine} + \text{Ret}_{325} + C_2S + C_3A + C_4AF + EA$$

Ecuación N.10. Ecuación Modelo de regresión lineal elegido

$$R28 = 31.190 - 0.002 - 9.901 - 0.167 + 1.805 + 2.766 - 6.789$$

5.4 Validar el modelo de regresión lineal planteado.

5.4.1 Independencia de los residuos

El test de Durbin- Watson como se refirió en el marco teórico informa sobre el grado de independencia entre los residuos, puede asumir valores entre 0 y 4. Sin embargo, si tiene un valor entre 1,5 y 2,5, se asume independencia. Realizando la estimación de este estadístico en la herramienta Rstudio, se obtiene lo siguiente:

```

Durbin-watson test
data: modelo4
Dw = 1.7061, p-value = 0.01731
alternative hypothesis: true autocorrelation is greater than 0

```

Figura N.27. Resultado del test Durbin – Watson sobre el modelo en RStudio

El valor arrojado por el test Durbin-Watson es de 1.7061, de tal manera que se puede afirmar la existencia de independencia entre los residuos.

Otra forma de visualizar la independencia de los residuos es graficándolos frente a cada variable y visualizar si existe algún patrón de tendencia, a continuación se ilustran los gráficos:

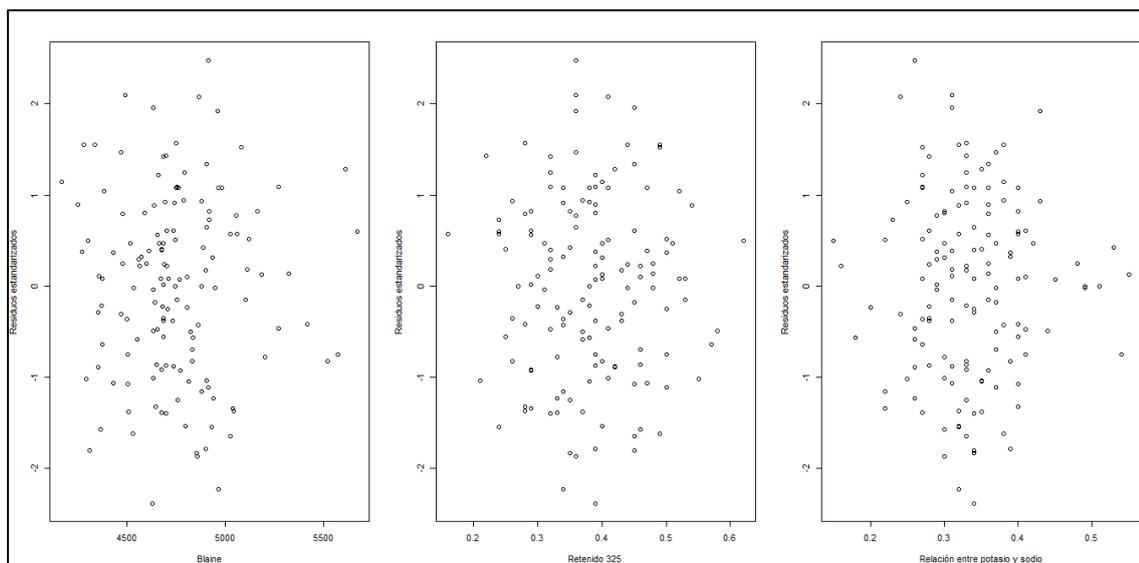


Figura N.28. Diagrama de residuos del modelo frente a las variables Blaine, retenido 325, relación entre potasio y sodio.

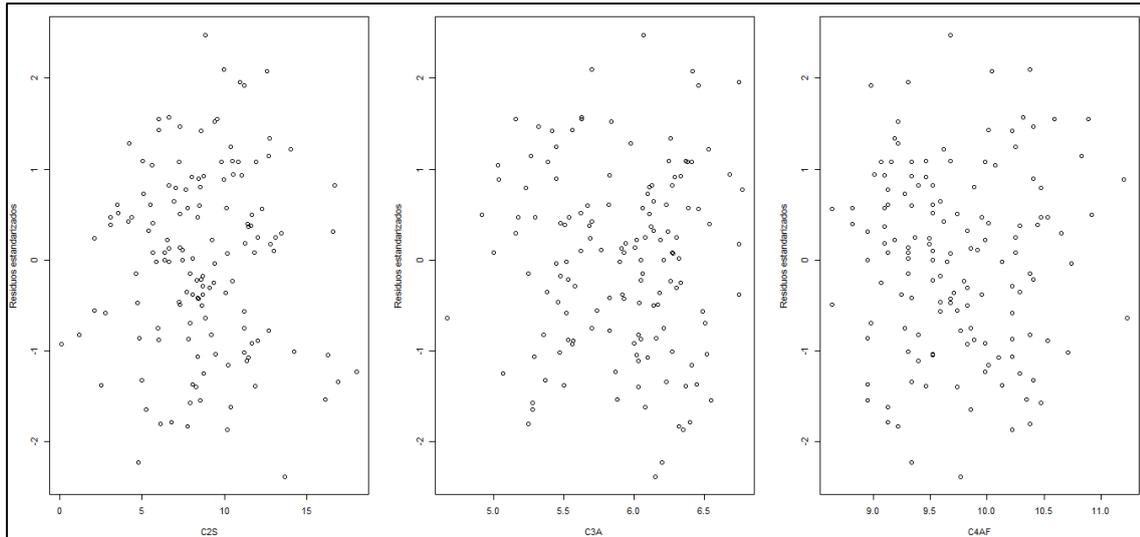


Figura N.29. Diagrama de residuos del modelo frente a las variables C2S, C3A, C4AF

Como se puede ver en las gráficas no hay existencia de alguna tendencia entre las variables analizadas y los residuos del modelo.

5.4.2 Normalidad de los residuos

Realizando el test de Kolmogorov - Smirnov para estimar la normalidad de los residuos, se utiliza nuevamente la herramienta Rstudio arrojando lo siguiente:

```

Lilliefors (Kolmogorov-Smirnov) normality test
data: residuals(modelo4)
D = 0.047383, p-value = 0.6126

```

Figura N.30. Resultado en RStudio normalidad de los residuos

Nota: cabe aclarar que se utilizó la modificación del test de Kolmogorov-Smirnov conocida como test Lilliefors. El test Lilliefors asume que la media y varianza son desconocidas, estando especialmente desarrollado para contrastar la normalidad.

Para realizar la evaluación de lo que arroja la herramienta, es necesario plantear las siguientes

hipótesis:

H_0 : Los residuos tienen una distribución normal

H_1 : Los residuos no tienen una distribución normal

Para una incertidumbre de $\alpha=0.05$, se tiene como regla de decisión:

- Valor- $p \leq \alpha$; Se rechaza la hipótesis nula.
- Valor- $p > \alpha$; Se acepta la hipótesis nula.

Por lo anterior se concluye que los residuos siguen una distribución normal.

Otra forma de visualizar la normalidad es revisando gráficamente los residuos del modelo:

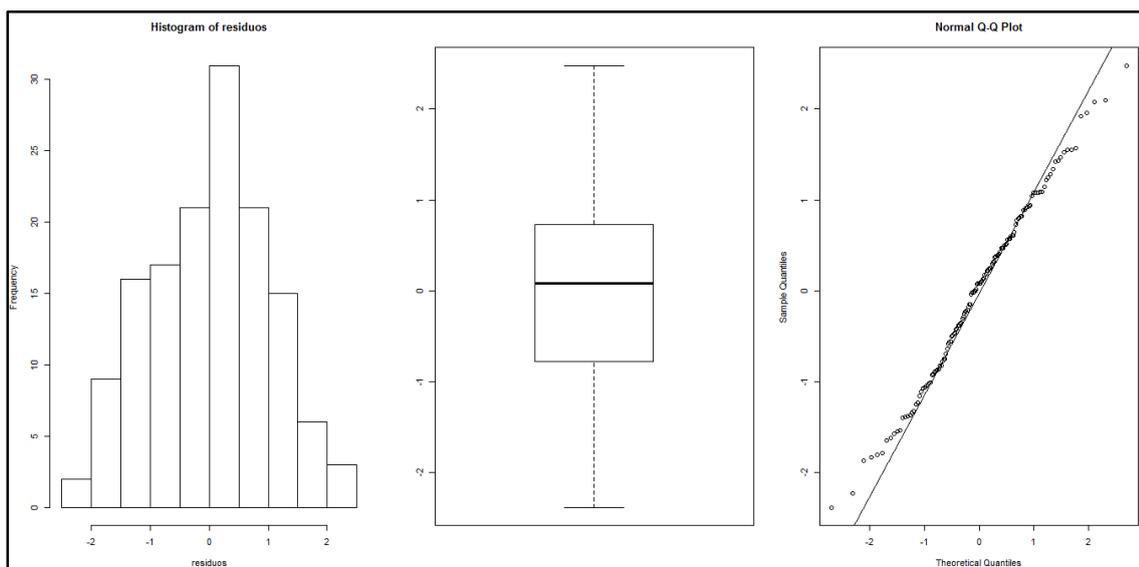


Figura N.31. Histograma, diagrama de caja y bigotes y gráfica Q-Q en RStudio para los residuos del modelo

Estos gráficos muestran un comportamiento de los residuos con tendencia de una distribución normal, confirmando lo obtenido por el test de Kolmogorov-Smirnov

5.4.3 Igualdad de varianzas

Para evaluar la igualdad de varianzas, se emplean los siguientes dos métodos:

- Método gráfico: en él se relacionan los residuos estandarizados frente a los valores ajustados. Con éste método se busca descartar alguna tendencia de los residuos frente a los valores ajustados entregados por el modelo.

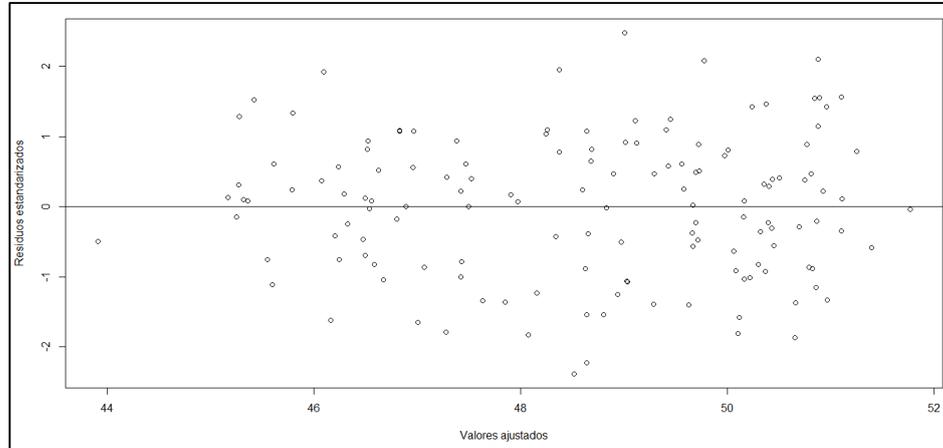


Figura N.32. Gráfico igualdad de varianzas

Visualmente no se evidencia una tendencia en particular.

- Empleando el test de igualdad de varianzas utilizando Rstudio se obtiene lo siguiente:

```

F test to compare two variances

data:  pp and qq
F = 3.3913, num df = 140, denom df = 140, p-value = 2.466e-12
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 2.431629 4.729673
sample estimates:
ratio of variances
 3.391285

```

Figura N.33. Resultados en RStudio para igualdad de varianzas

Para realizar la evaluación de lo que arroja la herramienta, se plantean las siguientes hipótesis:

H_0 : Las varianzas de los residuos son iguales

H_1 : Las varianzas de los residuos no son iguales

Para una incertidumbre de $\alpha=0.05$, se tiene como regla de decisión:

- Valor- $p \leq \alpha$; Se rechaza la hipótesis nula.

- Valor- $p > \alpha$; Se acepta la hipótesis nula.

Por lo anterior se concluye que las varianzas de los residuos no son iguales puesto que el valor- p es $2.466e-12$.

5.4.4 Multicolinealidad

Dentro de las herramientas más usadas para detectar multicolinealidad se encuentra:

- El factor de inflación de varianza, cabe aclarar que no se trata de un test estadístico que muestre si la existencia de multicolinealidad es grave, sino de reglas de decisión que tratan de establecer umbrales a partir de los cuales el problema planteado se considera preocupante. Tradicionalmente, valores del FIV superiores a 10 indican que el modelo de regresión lineal presenta un grado de multicolinealidad preocupante. Nuevamente se emplea la herramienta Rstudio, arrojando los siguientes valores:

Blaine	Ret_325	C2S	C3A	C4AF	EA
1.486120	1.125497	1.183069	2.729859	3.093341	1.121019

Figura N.34. Resultados RStudio multicolinealidad

Con lo cual se puede decir que la multicolinealidad de las variables no es preocupante.

- El método gráfico emplea las gráficas de dispersión que a continuación se presentan:

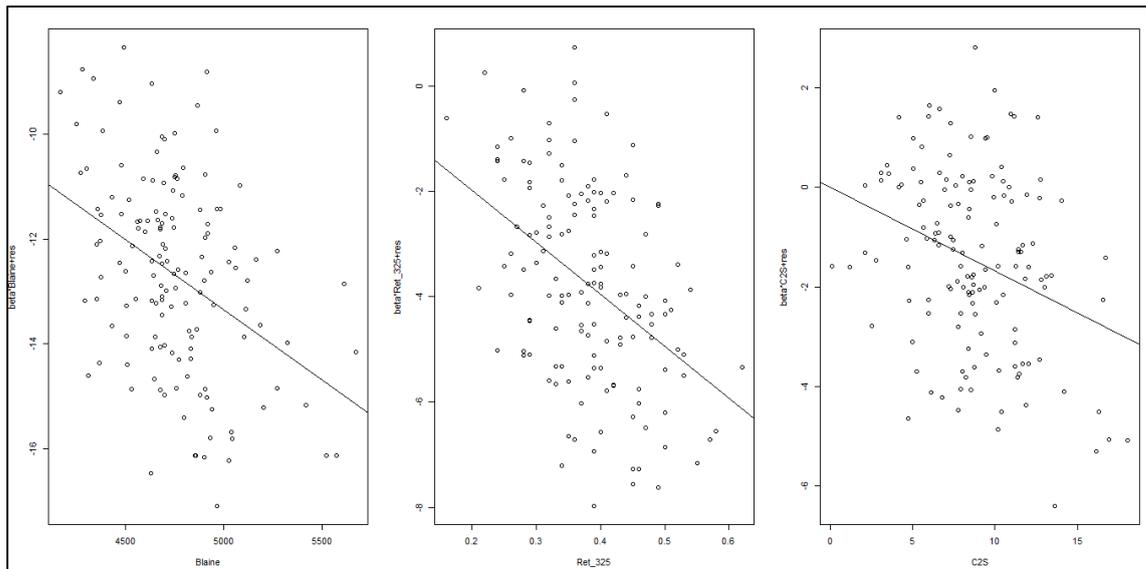


Figura N.35. Gráfico de dispersión para análisis de multicolinealidad Blaine, Ret 325, C3S

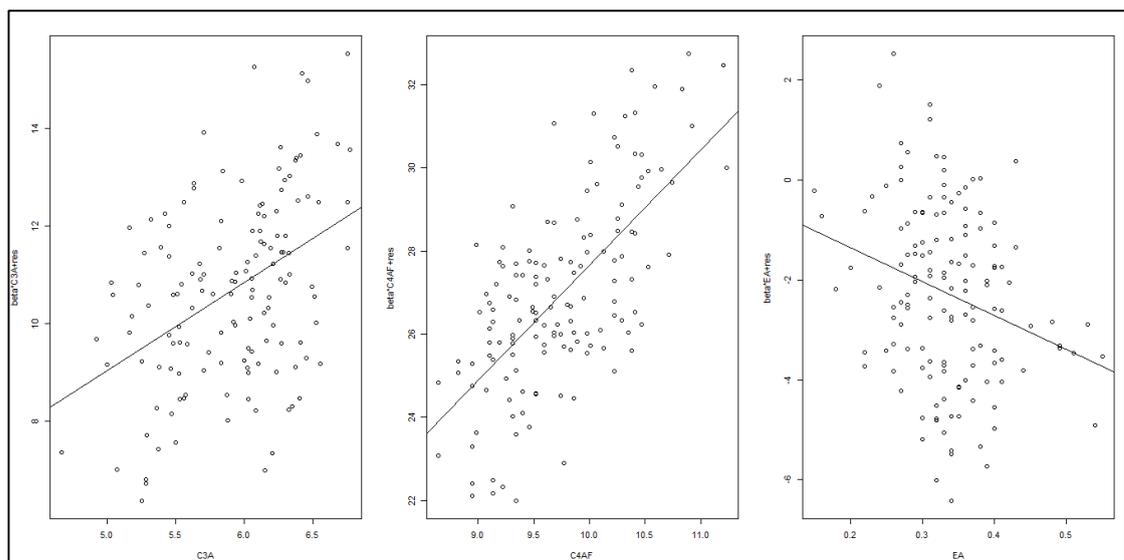


Figura N.36. Gráfico de dispersión para análisis de multicolinealidad C₃A, C₄AF, EA

Aunque los valores de colinealidad no son superiores a 10, sí se logra observar la diferencia en la tendencia de las variables químicas analizadas de C₃A y C₄AF, esto de acuerdo con los valores numéricos.

6. Conclusiones y Recomendaciones

Este proyecto parte de una necesidad fundamental en la industria cementera. Conocer un parámetro de calidad en uno de sus productos de mayor demanda “La resistencia del cemento Concretero a 28 días”. Es de destacar que son numerosas las investigaciones que han señalado relaciones entre las variables físicas en algunas de los materiales intermedios como el Clinker de cemento y la resistencia mecánica a la compresión del cemento. Por ejemplo, respecto a la consideración de las propiedades químicas del Clinker se busca conocer la relación con las variables físicas en el cemento por medio de regresión multivariada Valle García (2018), en “Un análisis estadístico de regresión lineal multivariado evidenció que el valor de la resistencia a la compresión del cemento sí es explicado por la concentración mineral en el clinker de alita, belita, ferrita y aluminato”. La resistencia mecánica es la característica de calidad más importante para el cemento Concretero. Esta característica solo puede ser medida una vez el producto está completamente terminado y su proceso de medición y seguimiento se realiza durante un periodo de 28 días. Por tal razón, resulta de gran importancia identificar métodos, modelos y ecuaciones para anticiparse su valor.

Los anteriores argumentos resaltan el papel relevante que tiene la predicción de las resistencias como parámetro de seguimiento al producto final, en el caso de este proyecto en el cemento concretero producido por un molino vertical, teniendo en cuenta esto la pregunta central de nuestra investigación fue:

¿Cuál es la relación existente entre las variables químicas y físicas y la resistencia del cemento Concretero a 28 días en un molino vertical?

Esto nos lleva a plantear los siguientes subproblemas:

¿Qué elaboraciones conceptuales y metodológicas sustentan investigaciones a nivel mundial para predecir la resistencia mecánica del cemento?

¿Hasta dónde han abordado los diferentes investigadores el problema de predecir la resistencia del cemento conociendo las variables físicas y químicas del cemento?

Para avanzar en la resolución del problema, se adoptó la metodología cuantitativa y descriptiva. La misma fue determinada, partiendo de los tres objetivos principales que la sustentan y de las metas que perseguimos que son: cuantificar, describir, explicar y validar los resultados. Esta metodología se adecua ya que la cuantificación y descripción surgen después de la exploración y tratamiento de datos, y sirve para organizar los resultados con el fin de que encajen con las explicaciones, y luego probar o validar las explicaciones. En sintonía, se elaboraron unas categorías de análisis y las siguientes dimensiones:

- a) Evaluación descriptiva de las variables que se miden regularmente en un laboratorio de control calidad en la industria cementera.
- b) Selección de las variables que afectan estadísticamente la resistencia del cemento concretero a 28 días.
- c) Creación del modelo predictivo para estimar la resistencia del cemento concretero.
- d) Evaluación del modelo propuesto a la luz de los supuestos para tener una mayor confiabilidad en las predicciones que se puedan construir con base a él.

De las cuales podemos determinar que las concepciones que se evaluaron en el modelo de regresión creemos que transmitirán a las siguientes investigaciones oportunidades de abordar otros aspectos que podrían ser evaluados en trabajos futuros, teniendo en cuenta las siguientes

premisas de acuerdo a los resultados:

- 1- Cuando se evaluó el supuesto de independencia de los residuos del modelo de regresión lineal, según el test de Durbin- Watson (Figura 27) los residuos si tienen independencia y no obedecen ningún patrón como se evidencia en los gráficos 28 y 29 para ser aceptado el modelo.
- 2- Evaluando la normalidad como segundo supuesto, se realizaron dos pruebas, la primera el test modificado de Kolmogorov - Smirnov conocido como test Lilliefors (Figura 30), el cual por evaluación de prueba de hipótesis confirma la hipótesis nula de aceptar que los residuos tienen una distribución normal. La segunda forma se evalúa la herramienta visual en la Figura 31 en donde se tiene la gráfica Q-Q plot en la que los valores graficados para los residuos obtenidos, a partir del modelo creado se apegan lo suficiente a la recta definida para esta prueba, evidenciando así, la normalidad de los residuos.
- 3- Revisando el tercer supuesto correspondiente al de igualdad de las varianzas de los residuos o más comúnmente conocido como homocedasticidad, se trabajó de dos maneras. En la primera por un método gráfico en la que se colocaron los residuos frente a los valores ajustados (Figura 32). Es deseable que los residuos aparezcan representados en una banda horizontal sin tendencias alrededor del cero. Lo cual fue evidenciado para el modelo. La segunda forma fue utilizando la herramienta Rstudio en el que tiene una función para evaluar homocedasticidad (Figura 33), la cual por una prueba de hipótesis arrojó que el modelo no tiene igualdad de varianza en sus residuos. Este problema de heterocedasticidad como se le conoce, aparece generalmente cuando

el modelo está mal especificado, bien sea en la relación de la respuesta con los predictores o bien en la distribución de la respuesta, bien en ambas cuestiones. En cualquiera de los casos se sugiere en trabajos futuros trabajar con estrategias como la Regresión Ridge, o técnicas de selección o disminución de variables por componentes principales.

- 4- Por último, se evaluó la existencia de multicolinealidad en el modelo de dos maneras: con el factor de inflación de la varianza (FIV) en el cual las variables con mayor valor fueron el C_3A y el C_4AF pero siempre menor a 10; con lo que se puede descartar multicolinealidad (Figura 34). La otra herramienta fue la visual en la que se graficaron cada variable frente al modelo en las Figuras 35 y 36, las cuales evidencian una asociación diferente en las variables ya mencionadas C_3A y el C_4AF ; pero que definitivamente no incumplen el supuesto.

En definitiva, se prevee la necesidad de seguir trabajando y utilizando herramientas estadísticas que permitan tener un mayor acercamiento al valor real de las resistencias a 28 días del cemento concretero. Un trabajo futuro puede ser la revisión de las variables influyentes en el cemento concretero a la luz de un análisis por componentes principales.

Bibliografía

- Aramburo Gonzalez, L. (2012). DISEÑO DE HERRAMIENTA PARA LA PROYECCIÓN Y CONTROL DE LA RESISTENCIA A LA COMPRESIÓN DEL CEMENTO CONCRETERO CON APLICACIÓN EN UNA EMPRESA CEMENTERA. ICESI.
- Barnett, V., & Lewis, T. (1994). OUTLIERS IN STATISTICAL DATA. (J. Wiley & Sons, Eds.) (3rd editio).
- Beckman, R. J., & Cook, R. D. (1983). OUTLIER... .. S. *Technometrics*, 25(2), 119–149.
- Cebrian, J. L., & Pisonero, F. (1971). DETERMINACIÓN DE SUPERFICIE ESPECÍFICA POR EL MÉTODO DE BLAINE, EN CENIZAS VOLANTES Y CEMENTOS PUZOLÁNICOS. *Materiales de Construcción*, 21, 81–91.
- Cortes Gomez, E. A., & Perilla Sastoque, J. E. (2014). ESTUDIO COMPARATIVO DE LAS CARACTERISTICAS FISICO-MECANICAS DE CUATRO CEMENTOS COMERCIALES PORTLAND TIPO I. Universidad Militar Nueva Granada.
- CPI. (2003). MOLIENDA DEL CEMENTO. *Cement Performance International*.
- E-302-07, I. (2008). FINURA DEL CEMENTO PORTLAND METODO DEL APARATO BLAINE (I.N.V.E-302-07). *Instituto Nacional de Vías*, 39(5), 561–563.
- Gnanadesikan, R., & Kettenring, J. R. (1972). ROBUST ESTIMATES, RESIDUALS AND OUTLIER DETECTION WITH MULTIRESPONSE DATA. *Biometrics*, 28(Special Multivariate), 81–124.
- Gorgas García, J., Cardiel López, N., & Zamorano Calvo, J. (2011). ESTADÍSTICA BÁSICA PARA ESTUDIANTES DE CIENCIAS. Madrid.
- Gutierrez De Lopez, L. (2003). EL CONCRETO Y OTROS MATERIALES PARA LA

- CONSTRUCCIÓN (Primera ed). Universidad Nacional De Colombia Sede Manizales.
- Hawkins, D. M. (1980). IDENTIFICATION OF OUTLIERS.
- Hidalgo de Cisneros Alonso, J. M. (2003). LOS MOLINOS VERTICALES EN LA INDUSTRIA DEL CEMENTO. *Loesche Española de Ingeniería, S. L.*, 191–193.
- Levine, David, M., Krehbiel, Timothy, C., & Mark L. (2012). ESTADÍSTICA DESCRIPTIVA (Primera ed). PEARSON.
- Matsumoto, S., Kamei, Y., Monden, A., & Matsumoto, K. (2008). COMPARISON OF OUTLIER DETECTION METHODS IN FAULT-PRONENESS MODELS, 49(3), 61–63.
- Montero Granados, R. (2016). MODELOS DE REGRESIÓN LINEAL MÚLTIPLE.
- Neville, A. M. (2013). TECNOLOGÍA DEL CONCRETO. M. en A. Soledad Moliné Venanzi.
- Otero, A., Ramirez, B., Bukasa, M., & Seke, A. (2016). EFECTO DE LA ADICIÓN DE CALIZA A CLÍNQUER PORTLAND ANGOLANO EN EL COMPORTAMIENTO MECÁNICO DE MORTEROS., (Noviembre).
- Quaglino, M., & Morello, J. (2012). MÉTODOS MULTIVARIADOS EN ESTUDIOS DE VULNERABILIDAD SOCIAL EN LA PROVINCIA DE SANTA FE.
- Rendón Macías, M. E., Villasís Keever, M. Á., & Miranda Novales, M. G. (2016). ESTADÍSTICA DESCRIPTIVA. *Revista Alergia México*, 63(4), 397.
- Rojo Abuín, J. M. (2007). REGRESIÓN LINEAL MÚLTIPLE. *Instituto de Economía y Geografía*.
- Ruíz Muñoz, D. (2004). MANUAL DE ESTADÍSTICA. *Eumed.net*.
- Rustom Jabbaz, A. (2012). ESTADÍSTICA DESCRIPTIVA, PROBABILIDAD E

INFERENCIA.

- Stutzman, P., Heckert, A., Tebbe, A., & Leigh, S. (n.d.). UNCERTAINTY IN BOGUE-CALCULATED PHASE COMPOSITION OF HYDRAULIC CEMENTS, 1–16.
- Tiwari, K., Mehta, K., Jain, N., Tiwari, R., & Kanda, G. (2007). SELECTING THE APPROPRIATE OUTLIER TREATMENT FOR COMMON INDUSTRY APPLICATIONS. *NESUG Statistics and Data Analysis*, 1–5.
- Valle García, D. E. (2018). MODELO DE REGRESIÓN LINEAL MULTIVARIADO APLICADO A LA RELACIÓN ENTRE VARIABLES QUÍMICAS DEL CLINKER Y VARIABLES FÍSICAS DEL CEMENTO. San Carlos De Guatemala.
- Walpole, R. E., Myers, R. H., & Myers, S. L. (1999). PROBABILIDAD Y ESTADÍSTICA PARA INGENIEROS (Sexta edic).

Anexos

Anexo 1. Aplicación de las ecuaciones de Bogue al cemento producido.

Muestra	C3S	C2S	C3A	C4AF	Muestra	C3S	C2S	C3A	C4AF
1	63.98	7.44	5.77	9.92	40	63.39	7.71	5.38	10.29
2	59.35	11.59	5.68	10.29	41	62.28	8.55	5.42	10.22
3	61.16	9.97	5.7	10.38	42	59.93	10.38	5.45	10.25
4	58.74	12.03	5.57	10.53	43	58.89	11.68	6	9.98
5	60.31	11.21	5.47	10.71	44	62.46	8.33	6.05	9.52
6	59.54	11.68	4.92	10.92	45	68.58	2.11	5.69	9.49
7	61.59	8.81	4.67	11.23	46	68.28	2.76	5.52	10.22
8	60.29	9.94	5.04	11.2	47	68.78	3.08	5.51	10.44
9	61.43	9.39	5.45	10.74	48	72.32	0.11	5.56	9.83
10	64.43	6.64	5.63	10.32	49	69.78	2.12	5.74	9.74
11	65.68	5.98	5.56	10.01	50	61.42	8.54	6.55	8.95
12	65.61	5.98	5.53	9.89	51	61.55	9.47	6.52	9.52
13	63.05	8.74	5.07	10.29	52	62.79	7.76	6.39	9.1
14	66.03	5.87	5.52	10.13	53	62.18	8.05	6.45	8.95
15	64.43	7.3	5.65	10.04	54	59.94	9.8	6.41	9.07
16	65.41	6.1	5.25	10.38	55	58.95	11.18	6.46	8.98
17	63.27	7.92	5.28	10.47	56	56.64	12.74	6.26	9.19
18	66.41	5.64	5	10.25	57	62.79	7.31	6.01	9.31
19	66.19	5.57	5.03	10.07	58	61.61	8.46	5.67	9.34
20	57.16	13.45	5.16	10.65	59	62.34	8.36	5.83	9.34
21	63.47	8.37	5.29	10.22	60	61	9.17	6.03	9.4
22	63.46	8.43	5.45	10.41	61	65.64	4.18	5.98	9.22
23	61.89	9.53	5.63	10.89	62	63.62	5.96	6.21	9.28
24	62.72	8.59	5.53	10.41	63	61.81	7.9	6.06	9.31
25	63.31	7.29	5.32	10.41	64	58.14	11.35	6.03	9.4
26	62.71	8.69	5.58	10.22	65	57	13.02	5.62	9.52
27	66.53	4.97	5.37	10.41	66	59.96	11.19	5.7	9.86
28	65.97	5.65	5.48	10.01	67	59.19	12.03	6.08	9.37
29	68.03	3.07	5.54	9.95	68	59.91	11.43	6.12	9.1
30	69.24	3.87	6.36	9.92	69	60.58	10.38	6.08	9.13
31	64.4	5.75	5.39	10.32	70	59.28	11.81	6.27	9.13
32	62.5	8.38	5.3	10.53	71	56.14	14.21	6.27	9.31
33	58.32	12.68	5.27	10.83	72	58.82	11.93	6.38	9.16
34	65.29	6.02	5.16	10.59	73	66.59	4.81	6.16	8.95
35	67.3	4.34	5.18	10.47	74	56.9	14.55	6.65	9.16
36	64.57	7.03	5.23	10.47	75	63.51	7.31	6.17	8.64
37	67.29	4.63	5.25	10.38	76	59.74	11.41	6.54	8.82
38	69.8	2.51	5.5	10.13	77	58.7	12.28	6.46	8.64
39	70.96	1.17	5.36	10.38	78	60.2	10.46	6.2	8.76
					79	64.66	7.44	6.21	8.95
					80	60.38	10.1	6.06	8.82

Muestra	C3S	C2S	C3A	C4AF	Muestra	C3S	C2S	C3A	C4AF
81	53.68	16.3	6.02	9.52	116	60.19	10.24	6.41	10.01
82	53.96	16.58	6.24	8.95	117	63.43	7.31	6.11	9.74
83	60.08	11.24	5.94	9.1	118	62.13	10.07	6.18	9.71
84	61.86	9.41	5.84	9.22	119	62.2	9.07	6.3	9.83
85	63.19	7.75	5.76	9.22	120	62.1	9.23	6.19	9.98
86	59.75	10.86	5.39	9.98	121	63.18	7.78	6.05	10.22
87	63.36	7.22	5.46	9.59	122	62.59	8.06	5.92	9.95
88	57.48	12.69	5.83	9.77	123	62.33	8.54	6.11	9.89
89	52.19	18.05	5.87	9.98	124	65.84	5.03	6.37	9.46
90	54.18	16.7	6.13	9.52	125	63.36	6.93	6.14	9.59
91	54.09	16.9	6.23	9.34	126	64.07	6.34	6.05	9.52
92	54.33	16.15	5.88	10.35	127	66.39	4.16	5.7	9.62
93	56.66	13.65	6.15	9.77	128	63.19	6.6	5.91	9.86
94	59.54	11.88	6.37	9.46	129	63.22	6.61	5.9	9.65
95	56.74	14.05	6.53	9.62	130	61.53	8.63	6.14	9.83
96	62.51	7.63	6.77	9.13	131	58.31	11.46	6.1	10.1
97	61.18	8.66	6.75	9.25	132	56.81	13.11	6.3	10.25
98	58.13	10.96	6.75	9.31	133	60.75	10.16	6.28	9.68
99	59.8	9.28	6.85	9.04	134	61.75	8.44	5.93	9.68
100	58.89	10.53	6.68	9.01	135	62.37	8.25	6.03	9.74
101	62.3	7.91	6.51	8.98	136	62.51	8	6.29	9.46
102	62.94	7.77	6.32	9.22	137	63.7	7.25	6.41	9.34
103	64.11	6.77	6.4	9.13	138	64.71	5.51	6.23	9.52
104	58.34	11.03	5.83	9.1	139	65.37	4.73	6.04	9.68
105	61.17	8.67	5.48	9.59	140	65.15	5.38	6.14	9.83
106	65.21	5.25	5.28	9.86	141	63.93	6.62	6.27	9.4
107	67.56	3.56	5.62	9.52	142	66.94	3.46	5.82	9.13
108	62.05	8.81	6.07	9.68	143	64.32	6.38	5.93	9.31
109	60.25	10.48	6.25	9.68	144	63.94	6.5	6.02	9.19
110	48.42	20.84	6.53	10.38	145	61.53	9.34	6.33	9.07
111	60.99	10.18	6.35	10.22	146	62.07	8.71	6.33	9.34
112	60.04	10.5	6.26	9.8	147	62.91	8.05	6.32	9.31
113	58.24	12.8	6.75	9.49	148	65.05	5.06	6.1	9.28
114	59.65	11.22	6.49	9.59	149	65.97	4.76	6.2	9.34
115	57.78	12.58	6.42	10.04					