

**CORPORACIÓN MEXICANA DE INVESTIGACIÓN EN
MATERIALES**

DIVISIÓN DE ESTUDIOS DE POSGRADO



**Modelo de Regresión Cuantil para analizar la resistencia al
desgaste de un recubrimiento depositado sobre Acero H13.**

TESIS

**Que para obtener el grado académico de
Maestra en Ciencia y Tecnología en INGENIERÍA INDUSTRIAL
Y DE MANUFACTURA**

POR

Darelys de la Caridad Bernal Iznaga

SALTILLO, COAHUILA.

14 de Diciembre de 2018

Modelo de Regresión Cuantil para analizar la
resistencia al desgaste de un recubrimiento
depositado sobre Acero H13.

Por

Darelys de la Caridad Bernal Iznaga

TESIS

Presentada al programa Interinstitucional en Ciencia
y Tecnología

Sede

Corporación Mexicana de Investigación en Materiales S.A. de C.V.

Como requisito parcial para obtener el grado académico de

Maestra en Ciencia y Tecnología

en Ingeniería Industrial y de Manufactura

**Programa Interinstitucional en Ciencia y
Tecnología COMIMSA**

Saltillo Coahuila, 14 de Diciembre de 2018

Corporación Mexicana de Investigación en Materiales, S.A de C.V
Gerencia de Desarrollo Humano
División de Estudios de Posgrado

Los abajo firmantes, miembros del Comité Tutorial recomendamos que la Tesis **“MODELO DE REGRESIÓN CUANTIL PARA ANALIZAR LA RESISTENCIA AL DESGASTE DE UN RECUBRIMIENTO DEPOSITADO SOBRE ACERO H13”**, realizada por la alumna **DARELYS DE LA CARIDAD BERNAL IZNAGA** matrícula **1706IM1127** sea aceptada para su defensa como Maestra en Ciencia y Tecnología en **INGENIERÍA INDUSTRIAL Y DE MANUFACTURA**.

El Comité Tutorial

Tutor Académico

Director de Tesis

Dr. David S. González González

Asesor

Dra. Argelia F. Miranda Pérez

Asesor

Dr. Rolando J. Praga Alejo

Vo. Bo

Dr. Pedro Pérez Villanueva

Coordinación General de Estudios de Posgrado

COMIMSA

Corporación Mexicana de Investigación en Materiales
Gerencia de Desarrollo Humano
División de Estudios de Posgrado

Los abajo firmantes, miembros del Jurado del Examen de Grado de la alumna, **DARELYS DE LA CARIDAD BERNAL IZNAGA**, una vez leída y revisada la tesis titulada **“MODELO DE REGRESIÓN CUANTIL PARA ANALIZAR LA RESISTENCIA AL DESGASTE DE UN RECUBRIMIENTO DEPOSITADO SOBRE ACERO H13”**, aceptamos que la referida tesis revisada y corregida sea presentada por la alumna para aspirar al grado de Maestra en Ciencia y Tecnología en **INGENIERÍA INDUSTRIAL Y DE MANUFACTURA** durante el Examen de Grado correspondiente.

Y para que así conste firmamos la presente a los 14 días del mes de Diciembre del año 2018.

Dr. Bernardo González Ortíz
Presidente

Dra. Indira G. Escamilla Salazar
Secretaria

Dr. David S. González González
Vocal

Dedicado

A mi mamá y a mi abuela

Por su inmenso amor y su apoyo incondicional

A mis hermanos y cuñados

por ser como padres para mí

A Eduardo y su mamá

por ser esa parte de mi familia aquí presente

Agradecimientos

A mi familia por su inmenso amor y por estar siempre para mí cuando más los he necesitado.

A mi tutor Dr. David González, por todo su tiempo y apoyo.

A mis asesores Dra. Argelia Miranda y Dr. Rolando Praga por compartir sus conocimientos.

A mis amigos, por los inolvidables momentos juntos.

A todos los trabajadores y alumnos de COMIMSA que me brindaron su ayuda durante el proyecto.

A los trabajadores de Soldatech, en especial al Msc. Alberto Valentín Zamora por haberme permitido realizar parte de mi proyecto allí y toda la ayuda y facilidades brindadas.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) y a la Corporación Mexicana de Investigación de Materiales (COMIMSA) por darme la oportunidad, así como a todos los académicos que de forma directa e indirecta aportaron a mi formación.

RESUMEN

En el presente trabajo se analiza la efectividad de un modelo de regresión cuantil, por ser una técnica de modelado de gran flexibilidad, para representar datos con alta variabilidad, reduciendo el efecto de valores atípicos. Con este fin, se analizará como caso de estudio, el fenómeno de desgaste en materiales, conocido por su comportamiento generalmente variable. Para ello se utilizó como respuesta la tasa de desgaste de acero H13 recubierto con Stellite 12, a través del proceso de Plasma Transferido por Arco (PTA). En base al análisis de la literatura, se seleccionaron los parámetros del PTA más influyentes en la calidad del recubrimiento, los cuales fueron utilizados como factores de entrada para la construcción del modelo. Para medir el desgaste se utilizó un indicador cuantitativo, conocido como tasa de desgaste.

El proceso de análisis se llevó a cabo en tres etapas. En la primera etapa se caracterizó y preparó el material, adicionalmente, se diseñó el experimento para ejecutar el depósito de los recubrimientos. La segunda etapa implicó la medición de dilución, microdureza y pruebas de desgaste, para finalmente en la tercera etapa, construir el modelo de regresión y ejecutar un algoritmo de optimización, para obtener los niveles de los parámetros del PTA que proporcionen la menor tasa de desgaste y a su vez un adecuado nivel de dilución.

Los resultados mostraron que a través de la regresión cuantil se puede cumplir con el objetivo de representar adecuadamente la tasa de desgaste del acero H13 recubierto mediante PTA. Por otro lado, a través del algoritmo multiobjetivo NSGA-II se lograron encontrar los parámetros óptimos para el proceso de PTA, de manera que se maximice la resistencia al desgaste del recubrimiento, y que a su vez se mantenga una dilución adecuada.

Palabras claves: Regresión cuantil, proceso de PTA, optimización NSGA-II.

Campo específico: Modelado y optimización de proceso.

TABLA DE CONTENIDO

Lista de Tablas.....	viii
Lista de Figuras.....	iv
Capítulo 1	
1. Introducción.....	1
Capítulo 2	
2. Planteamiento del Problema.....	4
2.1 Descripción del problema.....	4
2.2 Preguntas de Investigación.....	8
2.3 Hipótesis.....	8
2.4 Objetivo General.....	9
2.4.1 Objetivos Específicos	9
2.5 Justificación	10
2.6 Alcance y Delimitaciones	11
2.7 Aportaciones	11
Capítulo 3	
3. Estado del arte.....	12
3.1 Recubrimientos con PTA.....	18
3.2 Algoritmo de optimización multiobjetivo	21
Capítulo 4	
4. Marco Teórico	26
4.1 Acero H 13	26

4.2 Proceso de plasma transferido por arco (PTA)	27
4.3	Desgaste
.....	3030
4.3.1 Prueba de desgaste	31
4.3.2 Tasa de desgaste	32
4.4 Dilución	33
4.5 Diseño de Experimentos	34
4.5.1 Diseño Factorial	35
4.5.2 Diseño Central Compuesto	36
4.6 Regresión	37
4.6.1 Modelo de Regresión Lineal	38
4.6.2 Estimación de parámetros	39
4.6.3 Validación del modelo	44
4.7 Regresión Robusta	48
4.8 Regresión Cuantil	50
4.8.1 Cuantil	51
4.8.2 Modelo de Regresión Cuantil	54
4.8.3 Estimación	56
4.8.4 Inferencia en la regresión Cuantil	58
4.9 Optimización multiobjetivo	67
4.9.1 Algoritmo NSGA-II	70
Capítulo 5	
5. Metodología	75

5.1 Metodología Propuesta	75
---------------------------------	----

5.1.1 Pasos Generales de la Metodología Propuesta	75
Capítulo 6	
6. Aplicación	80
6.1 Determinación de las variables involucradas en el proceso (PTA) y niveles.....	81
6.2. Preparación de las muestras y diseño experimental preliminar	81
6.2.1 Caracterización y preparación del material	82
6.2.2 Material de aporte	83
6.3. Aplicación del recubrimiento mediante PTA	84
6.4. Pruebas de desgaste	85
6.5 Medición de la dilución	87
6.6 Construcción del modelo.....	91
6.7 Establecer las condiciones del proceso de PTA.....	94
Capítulo 7	
7. Conclusiones	97
Bibliografía	100

LISTA DE TABLAS

6.1	Párametros del PTA	80
6.2	Factor y dominio experimental	81
6.3	Composición química y dureza Acero H13	82
6.4	Composición química y propiedades físicas del Stellite	83
6.5	Matriz de experimentos preliminar.....	84
6.6	Condiciones del ensayo de desgaste	85
6.7	Resultados de la medición del desgaste y la dilución	87
6.8	Diseño de experimentos central compuesto	90
6.9	Coefficientes de ajuste para los modelos.....	91
6.10	Inferencia para el cuantil .99.....	94
6.11	Inferencia para el cuantil.01.....	94
6.12	Restricciones del problema de optimización.....	95
6.13	Resultados de la optimización.....	96
6.14	Resultados de la validación en campo	96

LISTA DE FIGURAS

4.1 Esquema de la antorcha de PTA.....	31
4.2 Aproximación del área desgastada	35
4.3 Ilustración de como estimar la pendiente de una funcion cuantil	56
4.4 Diagrama esquemático del mecanismo de promoción de individuos del NSGA.II.....	74
5.1 Metodología propuesta	79
6.1 Muestra de acero H13	85
6.2 Tribómetro Aton	88
6.3 Perfilómetro conectado al tribómetro	89
6.4 Análisis de la dilución y el desgaste	92

CAPÍTULO 1

1. INTRODUCCIÓN

Para decir que un proceso funciona satisfactoriamente es necesario analizar si sus variables de salida o de respuesta cumplen con ciertas y determinadas especificaciones. Evaluar su capacidad es analizar que tan bien sus variables de salida cumplen con las especificaciones. En tal sentido resulta útil recurrir al empleo de herramientas estadísticas como el diseño de experimentos, la regresión estadística, entre otras; de utilidad para extraer información de un proceso y analizar su comportamiento.

En el análisis de datos se considera la variabilidad (ruido), porque es un elemento necesario para darle sentido a los mismos. Phatak y Robinson (2005), consideran la variabilidad como ...” una situación en la cual las observaciones o las medidas deberían ser las mismas, pero no lo son”
....

Para Watson, Callingham y Kelly (2007) la comprensión de la variabilidad se relaciona con cambios previstos e imprevistos y con valores atípicos en las situaciones o en los datos. Todo esto influye en la decisión del tipo de modelo que se debe utilizar. Dado que la

variabilidad tiene consecuencias prácticas a causa de la dispersión, es difícil establecer modelos predictivos, para ello es necesario crear modelos estadísticos basados en el principio de erradicar el “ruido” creado por la dispersión.

Cuando se quiere controlar un proceso se necesita de un modelo que relacione entradas y salidas. El análisis de regresión lineal es una técnica para investigar y modelar la relación entre las variables, este método ayuda a comprender como un valor típico de una variable de respuesta cambia cuando alguna variable regresora (Factor) es modificada.

Existen modelos de regresión lineal, los cuales deben cumplir con una serie de supuestos, sin embargo, no siempre se cumplen, afectando de manera significativa la toma de decisiones acerca del proceso. Con este fin, existen procedimientos para construir modelos y hacer predicciones acertadas, incluso si alguno de los supuestos no se considera. La regresión robusta es un modelo insensible a valores atípicos, por lo tanto, es posible construir un modelo de regresión lineal sin considerar los supuestos tradicionales; como es el caso de la regresión cuantil, el cual se basa en la minimización de las desviaciones absolutas ponderadas con pesos asimétricos.

Resulta acertado entonces, emplear la regresión cuantil para modelar el fenómeno del desgaste; por su capacidad de robustez, dada por la

ausencia de supuestos, lo cual le permite representar datos con gran variabilidad. Por la literatura es conocido que el desgaste presenta un comportamiento generalmente variable, lo que lo ha hecho un fenómeno difícil de representar y controlar; convirtiéndose en centro de atención para los investigadores.

El presente trabajo pretende obtener como resultado un modelo que represente adecuadamente la variabilidad de este fenómeno, visto a través de un recubrimiento mediante PTA sobre piezas de acero H13; y que además pueda ser utilizado como función objetivo para optimizar y encontrar los niveles de los parámetros del proceso de PTA que minimicen el desgaste.

Este documento se estructura en 6 capítulos. En el capítulo 2 se describe el problema y las preguntas de investigación, la hipótesis, objetivos, justificación y el alcance del presente trabajo. En el capítulo 3 se realiza una revisión de algunos trabajos antecesores que abordaron temas relacionados con el problema planteado y se desarrolla el marco conceptual de la presente investigación en el cuarto capítulo. En el capítulo 5 se propone la metodología, reflejando los pasos a seguir para la elaboración del proyecto; por último, el capítulo 6 muestra la aplicación de la metodología a través de la experimentación, así como los resultados que se obtienen. Para finalizar se presentan las conclusiones y recomendaciones para trabajos futuros.

CAPÍTULO 2

2. Planteamiento del Problema

En este capítulo se presenta una descripción detallada del problema de investigación, a partir del cual se plantean las preguntas de investigación, hipótesis, objetivos, y justificación; así como el alcance de la presente.

2.1 Descripción del problema

Los modelos generalmente pueden clasificarse como paramétricos y no paramétricos, dependiendo de las suposiciones realizadas acerca de la distribución estadística de los datos en estudio (Harrell, 2001). Comúnmente, los modelos paramétricos son construidos a través de la distribución de los datos en cuestión, entre los que destacan: Exponencial, Weibull, Gamma y Log-Normal.

Cuando se trabajan con datos donde existe mucha variabilidad los modelos paramétricos suelen no ajustar, por lo que se vuelve necesario considerar métodos alternativos para modelar el comportamiento del fenómeno. Adicionalmente, es indispensable incorporar herramientas

estadísticas que permitan medir el ajuste de los modelos propuestos para seleccionar el que mejor represente los datos.

Un caso específico que con frecuencia se caracteriza por la variabilidad de sus datos es el fenómeno del desgaste en piezas y materiales. Por lo tanto, en ocasiones utilizar modelos distribucionales paramétricos para representarlo resulta una decisión poco acertada.

Fox y Weisberg, (2013) mostraron los inconvenientes de modelos convencionales (considerando los supuestos). Encontraron que los estimadores de mínimos cuadrados se ven afectados cuando la distribución de los errores no es normal, es decir la falta del cumplimiento de alguno de los supuestos del modelo de regresión lineal puede causar predicciones erróneas.

Debido a lo anterior, es de interés encontrar métodos que sean capaces de construir un modelo para representar y predecir de forma adecuada el tiempo de desgaste del material independientemente del cumplimiento de los supuestos del modelo de regresión.

En tal sentido sería útil recurrir a la regresión semi-paramétrica y no paramétrica. Al contrario de lo que sucede en la regresión lineal clásica, el objetivo no es encontrar un modelo único que describa y explique un fenómeno, sino obtener un método predictivo eficiente. Es

numéricamente intensiva, puesto que para cada observación se calcula un nuevo modelo.

Existen técnicas que son capaces de construir modelos de regresión sin considerar algún supuesto, es decir, modelos más robustos que permiten eludir algunas limitaciones tradicionales de los métodos paramétricos. Estos métodos están diseñados para no ser excesivamente afectados por violaciones de los supuestos, o por el proceso de generación de datos subyacente.

Un caso particular de la regresión robusta, lo constituye la regresión cuantil. Este método permite realizar la estimación de los diferentes percentiles (como, por ejemplo, la mediana) de una muestra de interés, ofreciendo la posibilidad de crear no solo una, sino distintas rectas de regresión para distintos cuantiles de la variable endógena a través de un método de estimación que se ve menos perjudicado en presencia de valores atípicos, permitiendo obtener un análisis más completo y robusto de los datos.

Otra de las ventajas de este tipo de regresión radica en la posibilidad de estimar cualquier cuantil, pudiendo valorar así lo que ocurre con valores extremos de la población. El método de regresión cuantil resulta de gran utilidad, ya que no utiliza los mínimos cuadrados para estimar los

parámetros del modelo y, por lo tanto, no requiere supuestos adicionales; además contempla la existencia de datos censurados y es capaz de construir un modelo que represente de forma adecuada el desgaste. Este modelo puede ser usado para realizar predicciones acertadas.

La regresión cuantil relaja los supuestos de la regresión clásica. Además, ofrece al investigador una visión, no alcanzable a través de la regresión clásica, del efecto de las variables explicativas sobre la localización (central y no central). La regresión cuantil (QR) permite estudiar la distribución condicional de Y en X en diferentes lugares y ofreciendo así una visión global de las interrelaciones entre ellas (Davino et.al., 2014). En particular, el fenómeno de desgaste presenta alta variabilidad, por lo cual no es recomendable el uso de modelos distribucionales paramétricos para modelar su comportamiento. Dado que es un fenómeno muy presente en los materiales presentes en máquinas y herramientas, resulta de gran importancia su estudio. Por ejemplo, el caso de la soldadura por fricción agitación, la herramienta se encuentra expuesta a un constante desgaste, por lo que se investiga la posibilidad de utilizar un material para su construcción que pueda ser tratado superficialmente de forma tal que se aumente su resistencia al desgaste, por ello, esta investigación se enfoca en el análisis y solución del problema que representa construir un modelo de regresión cuantil capaz

de representar y predecir el desgaste de un recubrimiento depositado mediante PTA sobre un acero H13, para aumentar su resistencia al desgaste, el cual representa el caso de estudio del presente trabajo.

2.2 Preguntas de Investigación

- 1.- ¿Cómo medir el desgaste?
- 2.- ¿Cuál sería la región experimental para las variables que intervienen en el tratamiento superficial PTA?
- 3.- ¿Cómo se puede a través del método de regresión por cuantiles establecer los niveles adecuados para mejorar la resistencia al desgaste del recubrimiento?
- 4.- ¿Qué método de estimación utilizar para obtener los parámetros del modelo?
- 5.- ¿Qué método de optimización es útil para establecer los niveles adecuados de los factores del proceso PTA dadas las características del modelo de regresión cuantil?

2.3 Hipótesis

Es posible construir un modelo de regresión cuantil, con el fin de modelar y aumentar la resistencia al desgaste del acero H13, teniendo

en cuenta las condiciones del recubrimiento superficial depositado mediante el proceso de PTA.

2.4 Objetivo General

Determinar un modelo de regresión por cuantiles que represente adecuadamente el comportamiento del desgaste de un recubrimiento depositado sobre acero H13 mediante PTA, para establecer las condiciones de proceso donde se mejore su resistencia.

2.4.1 Objetivos Específicos

1. Determinar las variables involucradas del proceso de PTA.
2. Identificar los factores del proceso PTA que tienen efecto significativo sobre el desgaste del acero H13 recubierto mediante PTA
3. Establecer los criterios que permitan organizar y diseñar las pruebas de desgaste.
4. Construir un modelo de regresión cuantil capaz de representar datos que manifiesten variabilidad.
5. Seleccionar el modelo de regresión cuantil más adecuado en términos de su valor cuantil y R^2 .

6. Establecer el método de optimización adecuado para determinar los niveles donde se maximice la resistencia al desgaste del recubrimiento, considerando como función objetivo el modelo de regresión cuantil.

2.5 Justificación

En muchos sectores de la industria, se emplean equipos y herramientas que se encuentran expuestos a un contante desgaste, llegando al punto donde ya no pueden ser utilizados, y deben ser reemplazados en su totalidad, o al menos uno de sus componentes. Es importante entonces considerar que el reemplazo de una herramienta o parte de la misma, conlleva a elevar los costos en que incurren las empresas, ya que, en muchos casos, además, estas herramientas o componentes son fabricados con materiales por lo regular costosos. Por lo tanto, considerar reemplazar la herramienta en períodos cortos de tiempo no es una solución rentable para la empresa. Bajo este contexto resulta provechoso disponer de un modelo de regresión capaz de representar un fenómeno tan variable y complejo como el desgaste, a través del cual se pueda controlar este problema, de forma tal que pueda ser empleado para maximizar la vida útil de las herramientas.

2.6 Alcance y Delimitaciones

Alcance:

Establecer un modelo de regresión cuantil capaz de predecir la resistencia al desgaste del material tratado superficialmente y encontrar las condiciones del proceso que alarguen su vida útil.

En este trabajo se realizarán las pruebas de desgaste, y los resultados serán considerados para ser utilizados como parte de un proyecto mayor.

2.7 Aportaciones

Científico:

Propuesta de una metodología para maximizar la resistencia al desgaste de un material tratado superficialmente mediante el proceso PTA.

Tecnológico:

Establecer las condiciones del proceso PTA que maximicen la resistencia al desgaste del material tratado superficialmente.

CAPÍTULO 3

3. Estado del Arte

Durante la revisión bibliográfica se pudo comprobar como muchos autores defienden el uso de la regresión por cuantiles, ya que las estimaciones generan modelos más robustos y por consiguiente más confiables.

Algunos modelos suponen distintas formas para el modelo paramétrico (Weibull, Exponencial, Ganma, etc.) planteando el modelo de regresión para los parámetros de las funciones. El modelo de riesgo proporcional de Cox es el más utilizado debido a su sencillez y facilidad para interpretar los coeficientes, existen numerosas variaciones de éste que lo han convertido en uno de los métodos más utilizados para analizar el efecto de los covariables sobre tiempos de vida. Sin embargo, Portnoy (2003) demostró que la regresión cuantil puede analizar los tiempos de vida para ofrecer un valioso complemento al modelo tradicional de riesgo proporcional de Cox. En su trabajo mostró la desventaja del enfoque de Cox, que fue obstaculizado por la falta de un estimador cuantil condicional para datos censurados, el cual es directamente análogo al estimador de Kaplan-Meier y se aplica bajo los supuestos

tradicionales para modelos de regresión censurados; Kaplan–Meier es un estimador no paramétrico que fue introducido por Edward L. Kaplan y Paul Meier en 1958 , y considera los datos censurados.

Portnoy proporcionó un algoritmo que es una generalización de la estimación de Kaplan-Meier y además comparte la propiedad óptima de los cuantiles de regresión cuando existe censura. El objetivo principal de este artículo fue comparar el modelo tradicional de Cox con el método propuesto a través de un enfoque cuantil, para estudiar como los cuantiles de regresión censurados complementaron el enfoque de Cox.

El autor utilizó un enfoque de regresión por cuantiles censurada, que es la generalización directa del estimador Kaplan-Meier de una muestra. Dicho enfoque, identifica fuentes de heterogeneidad que no pueden ser modeladas con métodos de riesgo-proporcional. Portnoy concluye que modelo de riesgos proporcionales de Cox, a pesar de que ha proporcionado un enfoque natural al análisis de datos de supervivencia, tiene significativas restricciones, sobre el comportamiento de las funciones de supervivencia condicional, pues la tasa de riesgo debe ser proporcional a lo largo del tiempo. Sin embargo, un enfoque de regresión cuantil puede proporcionar una flexibilidad sustancialmente mayor dentro del mismo paradigma fácilmente interpretable de modelado lineal.

En el 2006, Lee y colaboradores, mostraron la superioridad del análisis de la regresión cuantil sobre el método de regresión clásica. Según los autores, mediante la regresión cuantil se puede modelar cualquier número de cuantiles condicionales de una misma muestra, y se logra que los parámetros estimados para el modelo cumplan con la característica de ser insesgados. Sin embargo, según las dimensiones de la muestra, solo una cantidad determinada de cuantiles estimados podrán ser distintos.

Además, plantean que la estimación de parámetros en la regresión cuantil tiene la característica de ser robusta ante la existencia de observaciones extremas y heterocedasticidad. Esto se debe a la diferencia en la forma de realizar la estimación, a diferencia de la regresión clásica donde los parámetros se estiman minimizando la suma de cuadrados de residuos, en la regresión cuantil se estiman minimizando la suma de los valores absolutos ponderados de residuos.

Leng y Tong (2013), propusieron un estimador de regresión cuantil censurado motivado por ecuaciones de estimación no sesgada. Bajo el supuesto de independencia condicional del tiempo de supervivencia y el tiempo de censura de las covariables, demostrando que el estimador propuesto es consistente y asintóticamente normal. Desarrollaron un algoritmo computacional eficiente que utiliza el código de regresión

cuantil; la inferencia se realizó mediante el método de bootstrap. En conclusión, la regresión cuantil resultó un modelo más flexible que asume la independencia condicional del tiempo de supervivencia y el tiempo de censura. Además, propusieron un algoritmo nuevo y rápido, aplicable al modelo de regresión mediana de Ying.

Este procedimiento de inferencia basado en el muestreo se implementó de manera eficiente. Comprobaron su estimador con los enfoques desarrollados en Portnoy, entre otros, los resultados de la simulación mostraron que el nuevo método es útil y puede tener ciertas ventajas sobre los otros métodos, especialmente cuando se viola la linealidad global o la independencia incondicional de C y T (denotan T como el tiempo de vida y $C \leq T_0$ como el tiempo de censura, donde T_0 es el mayor tiempo de estudio de seguimiento) no se cumple.

Para datos de supervivencia en ensayos clínicos de oncología, Wu y Yin (2013) propusieron regresión cuantil bajo el sistema de censura común, los tiempos de supervivencia y los tiempos de censura son condicionalmente independientes dadas las covariables. En una formulación de mezcla, aplicaron la regresión cuantil censurada para modelar los tiempos de supervivencia de sujetos susceptibles y la

regresión logística para modelar los indicadores de si los pacientes son susceptibles.

Bajo la habitual censura independiente condicional, adoptaron el modelo de regresión cuantil de datos censurados por Peng y Huang (2008) para ajustar los tiempos de supervivencia de sujetos susceptibles y usar la regresión logística para modelar la fracción de cura. Tanto los cuantiles de regresión como los parámetros de la tasa de curación se calculan mediante ecuaciones de martingala (Fleming y Harrington 1991). Desarrollaron dos métodos de estimación: un enfoque utiliza completamente todos los cuantiles de regresión por iteración de la estimación entre la tasa de curación y los parámetros de regresión cuantil; y el otro separa a los dos a través de un estimador de suavizado no paramétrico de Kernel. Establecieron la coherencia uniforme y las propiedades de convergencia de la confianza para el estimador obtenido a partir de ambos métodos.

En el trabajo de Sun y colaboradores publicado en el 2016, consideraron el modelo de regresión cuantil para analizar datos longitudinales, dada su robustez en el manejo de resultados sesgados y censurados; modelo que además es flexible para acomodar las relaciones dinámicas covariable-resultado. Específicamente, estudiaron un modelo de regresión cuantil longitudinal que especifica los efectos de covariable

en los cuantiles marginales del resultado longitudinal. Este modelo permitió acomodar los cambios dinámicos de los perfiles de resultados con el tiempo. Propusieron un procedimiento de estimación e inferencia que consideró de forma apropiada la censura y el seguimiento irregular dependiente de los resultados, el cual puede implementarse fácilmente basándose en el software existente para la regresión cuantil. Establecieron las propiedades asintóticas del estimador propuesto, incluyendo consistencia uniforme y débil convergencia.

Como resultado obtuvieron un estimador ponderado de relación de intensidad inversa que pudo corregir eficazmente el sesgo debido al seguimiento irregular dependiente del resultado, con un modelado razonable del proceso de tiempo de seguimiento. Supusieron que los datos censurados por la izquierda son una constante fija, este nuevo método puede adaptarse fácilmente a los casos en que la variable censurada por la izquierda es aleatoria pero siempre observada. Además, es fácil extenderlo para hacer frente a resultados longitudinales doblemente censurados, por ejemplo, mediciones sujetas a un límite de detección superior, así como un límite de detección inferior.

Posteriormente Wu y Yin (2017), expusieron un enfoque de imputación múltiple para la regresión cuantil considerando datos censurados con una fracción de supervivencia. Desarrollaron un algoritmo iterativo para

estimar la probabilidad condicional para cada sujeto. Utilizaron esta probabilidad estimada y la imputación de la muestra de Bernoulli, para clasificar cada sujeto como curado o no curado y luego emplearon el método localmente ponderado para estimar los coeficientes de regresión cuantil con sólo los sujetos no curados. Al repetir el procedimiento de imputación varias veces y tomando un promedio sobre los estimadores resultantes, obtuvieron estimadores consistentes para los coeficientes de regresión cuantil. Su enfoque relajó la hipótesis de la linealidad global habitual, permitiendo aplicar la regresión cuantil a cualquier cuantil particular de interés. Establecieron propiedades asintóticas para los estimadores propuestos, incluyendo tanto la consistencia como la normalidad asintótica.

El análisis del Estado del Arte ha permitido conocer varios enfoques del método de regresión por cuantiles para la predicción de características de interés, los cuales han sido comparados con otros métodos tradicionales, demostrando así que los resultados obtenidos mediante la regresión por cuantiles tienen mayor representatividad y robustez.

3.1 Recubrimientos con PTA

Se han encontrado trabajos en la literatura referentes al proceso de PTA, considerándolo de gran utilidad en el recubrimiento de materiales para

aportar mayor resistencia al desgaste. Un caso es el trabajo presentado por (Zamora López, 2016), donde se estudia el efecto en la microestructura y la resistencia al desgaste de la adición de nanopartículas de carburo de tungsteno (WC) y carburo de titanio (TiC) en el metal de aporte en polvo Stellite ® 12; como material base se seleccionó el acero grado herramienta H13.

En este trabajo se realizó una variación del amperaje en 5 niveles, pues se determinó que influye en la dilución del recubrimiento depositado en el metal base. Los resultados evidenciaron que no existe mejora alguna en las de propiedades mecánicas del Stellite ® 12 con la adición de nanopartículas de TiC y WC según los porcentajes seleccionados. Esto se debe a que la dureza del recubrimiento se redujo en promedio 15 HRC con la adición de nanopartículas, respecto a la dureza del metal de aporte sin ser modificado (50 HRC), por lo que la presencia de nanopartículas no es necesaria. Por otra parte, de acuerdo al análisis de dilución, se determinó que las probetas que presentaron mejores características (entre 25% y 35% de dilución), fueron aquellas obtenidas con parámetro de 165 y 170 A de corriente de arco transferido, lo que para el autor resulta evidente ya que plantea que, entre mayor amperaje, mayor es la entrada de calor al sustrato y por consecuencia mayor penetración.

Aguirre (2012) presenta una investigación referente a la viabilidad de recuperación de componentes de acero grado herramienta D2 sometidos a desgaste, aplicando el proceso de Plasma Trasferido por Arco (PTA); utilizando como aporte en polvo una superaleación base níquel y tungsteno. En este trabajo se realizó una optimización de los parámetros del proceso en función del porcentaje de dilución, obteniendo como resultado un modelo con una predicción aceptable con respecto a los valores reales medidos durante la experimentación.

Para realizar la optimización, se aplicó la técnica de la función de deseabilidad, la cual se enfoca en una respuesta objetivo considerada durante la experimentación. El valor objetivo de 62.02 % fue sacado de la muestra que no presentó fractura y que sus características mecánicas y metalúrgicas fueron las mejores en el rango de parámetros evaluados, el valor inferior de 59.86 % se tomó como límite inferior de las muestras aceptables y el valor superior de 62.39 %, el cual se obtuvo del límite superior de las muestras que no fracturaron. Al aplicar el optimizador se obtuvo como resultado una dilución de 62.0139 %, con la cual, según el autor, la característica del cordón no debe presentar microfracturas; se generó como resultado un valor de corriente de 95.7137 A y velocidad 10.1994 cm/min, que fueron los factores variados durante la experimentación.

Molina (2016) determina el efecto de la adición de $WC_{(NPs)}$ en la microestructura y resistencia al desgaste de un recubrimiento base Ni obtenido mediante el proceso de plasma transferido por arco (PTA), depositado sobre superficies de acero ASTM A36. Los parámetros empleados para disminuir la dilución fueron encontrados en la literatura. Molina establece como parámetros para el proceso, aquellos más importantes que tienen relación directa en el porcentaje de dilución y por consiguiente en la calidad del recubrimiento, los cuales son: corriente del arco (150 A), velocidad de avance (17 cm/min), tasa de alimentación del polvo (20 g/min) y altura del arco (10 mm). Igualmente se establecieron los diferentes flujos de gas: gas de plasma y arrastre (5 l/min), y gas de protección (12 l/min). En este trabajo se varía el porcentaje de nano partículas de tungsteno adicionadas a la muestra.

3.2 Algoritmo de optimización multiobjetivo

Muchas veces los problemas reales deben ser resueltos teniendo en cuenta la existencia de múltiples medidas de actuación que deben ser optimizadas o satisfechas simultáneamente. Para ello se mezclan técnicas de computación evolutiva con la teoría de optimización multiobjetivo, obteniéndose los Algoritmos Evolutivos Multiobjetivo (MOEAs) que brindan la posibilidad de realizar búsquedas en espacios ilimitados y complejos.

Entre los algoritmos evolutivos más mencionados, se encuentra: NSGA, SPEA, NSGA-II, SPEA-II y PAES-II. Estos son considerados una herramienta algorítmica que permiten encontrarle la solución a varios problemas de optimización de objetivos múltiples, ya que a través de ellos se puede sustentar toda una población de soluciones óptimas (Alvarado et al.,2005).

Peñuela y Granada (2007) presentan conceptos básicos de la optimización multiobjetivo, enfocados hacia la descripción del algoritmo NSGA-II, Nondominated Sorting Genetic Algorithm por sus siglas en inglés, el cual es conocido en la literatura especializada por su buen desempeño y por los pocos cambios necesarios.

Loa autores antes mencionados presentan una herramienta que puede ser empleada en la solución de problemas de optimización multiobjetivo caracterizados por una determinada complejidad. Para ello realizan una explicación del algoritmo multiobjetivo en cuestión. El primer punto que destacan es la necesidad de conocer si el modelo de problema que se quiere optimizar, corresponde a un caso multiobjetivo o si simplemente se puede resolver utilizando una técnica mono-objetivo. En caso contrario consideran el algoritmo elitista NSGA-II, pues con un adecuado entendimiento del concepto de dominancia y de la manera en

que funcionan los algoritmos genéticos, resulta una herramienta fácil de comprender.

Peñuela y Granada plantean que este algoritmo tiene como característica el mantener las soluciones correspondientes al frente dominante dentro de los individuos de la población, a través de todas las generaciones del algoritmo. Además, debido a que los requerimientos desde el punto de vista computacional se consideran relativamente bajos, los mecanismos de diversidad se pueden implementar de una forma más rápida en comparación con otros algoritmos de su tipo, como el SPEA y DBGA.

También Correa y colaboradores (2008), presentan una filosofía del algoritmo multiobjetivo elitista NSGA-II, a través de su diagrama de funcionamiento y los engranajes especiales que posibilitan la conservación y desarrollo de soluciones Pareto-óptimas. Este algoritmo es empleado para resolver el “problema de la mochila”, conocido como un clásico de la optimización, acondicionado para optimización multiobjetivo. Desarrollan una propuesta alternativa para ser comparada con el esquema básico del NSGA-II.

Como resultado se percibe el funcionamiento que presenta el algoritmo para solucionar problemas de optimización multiobjetivo; al permitir la obtención de soluciones de adecuada calidad con su posterior desarrollo

hacia regiones Pareto-óptimas, gracias a su modelo de ordenamiento que permite conservar las mejores soluciones. El objetivo de Hernández y colaboradores (2016) fue la optimización multiobjetivo del proceso de laminación en caliente de láminas de acero en un tren acabador de seis pasos, implementando el algoritmo genético de búsqueda no dominada elitista de segunda generación NSGA-II con dos cruza diferentes (AG1 y AG2) y dos funciones objetivo. Resultando muy útil y necesario aplicar este algoritmo ya que, el problema de la programación de los pasos de laminación en caliente de aceros posee grandes espacios de búsqueda de soluciones, lo que aumenta su complejidad.

Se realizaron dos clases de experimentos computacionales, la búsqueda de la mejor solución y evaluación comparativa de las configuraciones del NSGA-II. Para cada experimento se obtuvieron 3000 soluciones. Se midió el tiempo de cómputo necesario para generar cada conjunto de soluciones y resultó que el AG1 tuvo menor tiempo de cómputo con respecto al AG2 y por lo tanto presentó el menor costo computacional. Para la experimentación se empleó una computadora con procesador Intel Core i7 con una velocidad máxima de procesamiento de 31 GHz y 12 GB RAM.

Como conclusión se obtuvo que la implementación del NSGA-II con sus variantes AG1 y AG2 minimizaron la flexión de los rodillos de

trabajo con respecto a la obtenida en datos industriales y mejoró en un 0.37 % el tiempo de laminación industrial. Se demostró, además que la aplicación de algoritmos genéticos multiobjetivo es factible para esta clase de problemas gracias a su fácil adaptación a modelos matemáticos complejos, con bajo costo computacional.

CAPÍTULO 4

Marco Teórico

En el siguiente capítulo se hace referencia a las herramientas que se desarrollaran en el transcurso del proyecto para ejecutar el modelado y estimación de la resistencia al desgaste del material acero H13.

4.1 Acero H 13

El acero grado herramienta H13, forma parte de los aceros para trabajos en caliente, comprendidos dentro del grupo H. Presenta una excelente combinación de dureza y resistencia a la fractura. Sus principales elementos aleantes son: cromo, molibdeno, vanadio, silicio y manganeso (Treviño, 2004).

Este tipo de acero se emplea en la fabricación de útiles o herramientas destinados a modificar la forma, tamaño y dimensiones de los materiales por cortadura, por presión o por arranque de viruta. Presentando una adecuada resistencia al desgaste, debido a su buena resistencia a la fatiga térmica. Resulta adecuado para temple al vacío y tratamientos superficiales como lo son la nitruración iónica, líquida o gaseosa y recubrimientos.

4.2 Proceso de plasma transferido por arco (PTA)

Una definición del proceso de plasma transferido por arco (PTA) lo señala como un proceso de soldadura de arco con gas protector, donde la coalescencia de los metales es lograda a través del calor de entrada transferido por un arco creado entre un electrodo de tungsteno y la pieza de trabajo (arco transferido) o entre el electrodo y la boquilla (arco no transferido).

También se emplea como tratamiento superficial en el cual se funde un polvo metálico en un arco de plasma, soldándose sobre la superficie de la pieza trabajada. El recubrimiento se forma por el polvo de metal, el cual pasa a través de una boquilla de cobre al plasma, la cual contiene en su interior un electrodo de tungsteno, y es depositado sobre el sustrato. El gas de plasma es conducido a través de esta boquilla, formándose un arco entre esta y la punta del electrodo, el cual es controlado a través de una fuente de poder que trabaja bajo ciertas frecuencias de oscilación.

El arco produce que el gas (generalmente se utiliza argón , helio o una combinación de ambos gases) se convierta en plasma e incremente su temperatura alcanzando niveles superiores a 10,000 Kelvin. A estos niveles de temperatura, el gas separa sus cargas negativas y se convierte

en iones positivos, generándose de esta manera lo que se conoce como plasma (Brett & Wolfe, 2010).

La fusión metalúrgica entre el sustrato y el recubrimiento se hace posible debido a que las partículas son calentadas por el plasma a temperaturas muy altas. El depósito formado es denso y resistente a la corrosión y el desgaste. Además, los recubrimientos se aplican con precisión, sobre una amplia gama de materiales de base, con cualquier configuración geométrica (Echevarría, 2016).

Debido a que el aporte de estos recubrimientos se caracteriza por ser en forma de polvo, hace que sean considerados como muy versátiles, ya produce una escala amplia de recuperaciones de componentes a través de recubrimientos, los cuales se caracterizan por tener porcentajes de dilución relativamente bajos (Aguirre Sánchez, 2012).

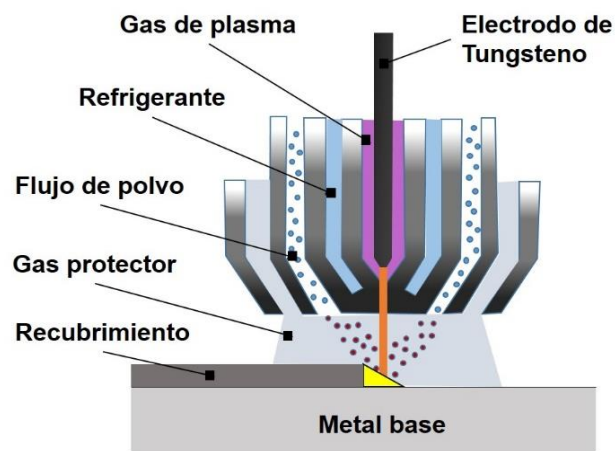


Figura 4.1. Esquema de la antorcha de PTA (Molina, 2016).

El grosor del depósito puede ser controlado alrededor de 0.2mm y 5mm en una sola capa, además de que este proceso puede ser automatizado, y además se pueden depositar varias capas sobre el material base.

Debido al alto grado de automatización el proceso PTA es ideal para piezas en serie y ofrece las siguientes ventajas:

- Alta reproducibilidad
- Bajo grado de dilución con el material base
- Zona afectada por el calor pequeña y concentrada
- Permite tasas elevadas de recubrimiento
- Gran variedad de aleaciones de polvos
- Combinaciones con materiales duros

Según Balasubramanian y colaboradores (2009) son varios los parámetros que intervienen en el proceso de PTA, pero según la incidencia que tienen en el porcentaje de dilución y la calidad del recubrimiento, se consideran más relevantes:

- Corriente del arco
- Velocidad de avance
- Tasa de alimentación del polvo
- Frecuencia de oscilación de la antorcha

- Altura del arco

4.3 Desgaste

Rabinowicz (1995) definió el desgaste como la erosión de material sufrida por una superficie sólida por acción de otra superficie, y además que está relacionado con las interacciones entre superficies y más específicamente con la eliminación de material de una superficie, como resultado de una acción mecánica.

El desgaste es el responsable más frecuente de modo de falla de los componentes de máquinas. Existen diversas clasificaciones para los diferentes tipos de desgaste. Se puede catalogar en dependencia del mecanismo que opera o en base a la existencia o escasez de lubricante. No existe unidad en los criterios reflejados en la literatura para determinar los tipos de desgaste que se pueden presentar.

A continuación se presentan tres categorías en las que de manera general se puede agrupar este fenómeno :

- Desgaste abrasivo
- Desgaste por deslizamiento
- Desgaste por erosión

4.3.1 Prueba de desgaste

Una prueba de vida acelerada es aquella en la cual un artículo o producto de interés, se somete a un esfuerzo en condiciones ambientales mayores a las que típicamente estará operando. Los principales objetivos de acelerar la vida de un producto son: estimar la distribución de vida de dicho producto, identificar fallas en el diseño, medir y demostrar la confiabilidad (Jiménez et al., 2009).

Las pruebas conocidas como “Pin-on-disk” tienen como objetivo someter el “pin” a condiciones severas de fricción para conocer el comportamiento de la pieza de interés al someterla a desgaste y obtener información de la mismas, en algunos casos los tiempos de vida.

El pin requerido para ejecutar la prueba de desgaste empleando este método tiene un tamaño pequeño, va colocado perpendicularmente sobre una pieza que usualmente es un disco redondo encima del cual se posiciona el material de estudio. La máquina de pruebas causa que el pin y el disco giren entre sí; como consecuencia de ello, se forma un camino de desgaste en el disco. El volumen del camino de desgaste o surco coincide con el material que se perdió en el tiempo en que se realizó la prueba, y a partir de él se puede calcular el coeficiente de desgaste.

Al realizar ensayos comparativos entre distintos materiales con las mismas condiciones, o el mismo material con distintas condiciones, se puede establecer una clasificación de materiales desde el punto de vista de resistencia al desgaste, pero no son valores absolutos (Aguirre Sánchez, 2012).

4.3.2 Tasa de desgaste

La tasa de desgaste se establece como un parámetro cuantitativo universal para el desgaste, ya que en ocasiones usar el volumen de material desgastado no es útil ni ilustrativo debido a las diferencias en las condiciones de prueba usadas, en general se utiliza la siguiente expresión (Luengo López, 2008) :

$$K = \frac{Vol}{W \times x} \left(\frac{mm^3}{N.m} \right) \quad (4.1)$$

Donde:

Vol : Volumen desgastado

W : Peso aplicado

x : Distancia recorrida.

Para obtener el volumen desgastado de la huella, se emplea la ecuación

(4.2):

$$Vol = A2\pi r \quad (4.2)$$

Donde

A: es el área promedio calculada.

r : radio de la huella desgastada

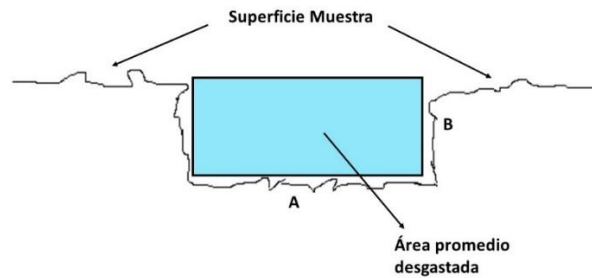


Figura 4.2. Aproximación del área desgastada (Luengo López, 2008)

La figura 4.2 representa la aproximación realizada para calcular el área de desgastada de la muestra.

4.4 Dilución

Ya en el recubrimiento duro por soldadura, el mayor interés sobre la geometría está enfocado sobre la dilución, como criterio que gobierna la calidad de los depósitos en cuanto a composición y consecuentemente en la durabilidad de los componentes bajo las condiciones de servicio (Cruz et al., 2011).

La dilución es el porcentaje de la unión soldada que corresponde al metal base, este valor se obtiene a través de la siguiente ecuación:

$$\% \text{ de dilución} = \frac{B}{A+B} * 100 \quad (4.3)$$

Donde

A: área de la sección transversal del metal depositado sobre la superficie original del metal base

B: área de la sección transversal de la zona de penetración en el metal base.

4.5 Diseño de experimentos

Como su nombre lo indica, la metodología del diseño de experimentos está basada en la experimentación. Según Montgomery (2005), consiste en realizar una serie de pruebas en las que se inducen cambios deliberados a las variables de un proceso para poder identificar las causas de los cambios en la respuesta de salida elegida. También se define como una técnica estructurada para poder eficientemente caracterizar, mejorar y optimizar un proceso o producto a través de la recolección, análisis e interpretación de los datos (Box et al. 2005).

El diseño de experimento tiene como objetivo estudiar si se produce alguna variación en un proceso, ya sea una mejora o no, cuando se utiliza

un determinado tratamiento sobre él. La metodología del diseño de experimentos permite obtener un conocimiento más amplio sobre el proceso en cuestión, al adentrarse en el análisis de cómo variar las condiciones habituales de realización de un proceso empírico, con el fin de incrementar la probabilidad de percibir cambios significativos en la respuesta. Entre los motivos por los que se hace necesario realizar un diseño de experimentos se encuentran el hecho de comparar las respuestas en diferentes niveles de observación de las variables controladas y obtener un modelo que permita hacer predicciones de respuestas futuras.

En ocasiones resulta útil desarrollar primeramente un “experimento piloto”, lo cual sería una experimentación preliminar en la que se emplean un número pequeño de observaciones. La realización de este permite complementar la lista de acciones a realizar y calcular la cantidad de observaciones que se requieren para el experimento principal, además de permitir detectar problemas no esperados durante la recogida de datos.

Existen varios tipos de diseños experimentales, dos de ellos son el diseño factorial y el diseño central compuesto.

4.5.1 Diseño Factorial

Los diseños factoriales 2^k son diseños en los que para cada factor k , se consideran solamente dos niveles (se suelen denotar + y -), se codifican dando el valor de +1 para el mayor y -1 para el menor. De esta manera se obtiene una matriz de diseño que considera todas las posibles combinaciones de los niveles de los k factores. Cada una de las filas representa un tratamiento.

El diseño 2^k requiere generalmente de un menor número de corridas, por lo que resulta muy útil utilizarlo como “experimentación piloto” en las primeras fases del trabajo experimental.

4.5.2 Diseño Central Compuesto

Los diseños centrales compuestos son diseños de tratamientos factoriales 2^k con $2k$ combinaciones adicionales las cuales son nombradas puntos axiales y n_c puntos centrales. Las coordenadas de los puntos axiales de los ejes del factor codificado son $(\pm\alpha, 0, 0, \dots, 0)$, $(0, \pm\alpha, 0, \dots, 0), \dots, (0, 0, 0, \dots, \pm\alpha)$, y los puntos centrales tienen la forma $(0, 0, 0, \dots, 0)$.

El diseño central compuesto tiene varias propiedades como ortogonalidad, rotabilidad y uniformidad, en consecuencia de la

selección de α en los puntos axiales. Pero solo la rotabilidad será valorada como una propiedad deseable en estos diseños. La rotabilidad significa que varianza de los valores estimados sea constante en puntos equidistantes del centro del diseño (Jiménez, 2015). El diseño se convierte en rotable según la elección de α , esta se calcula como:

$$\alpha = \left(2^k\right)^{\frac{1}{4}}.$$

4.6 Regresión

En el marco del análisis estadístico multidimensional interesa, en gran medida, descubrir la interdependencia o la relación existente entre dos o más de las características analizadas.

El análisis de la dependencia estadística admite dos planteamientos: el estudio del grado de dependencia existente entre las variables que queda recogido en la teoría de la correlación y la determinación de la estructura de dependencia que mejor exprese la relación, lo que es analizado a través de la regresión.

La regresión es una técnica estadística a través de la cual es posible modelar e investigar la relación entre variables. Sus aplicaciones son variadas en un amplio número de campos, incluyendo en ingeniería,

ciencias física y química, administración, entre otras, por lo que se adapta a una gran cantidad de situaciones. De hecho, puede ser que el análisis de regresión sea la técnica estadística más usada (Montgomery et al., 2006). En la investigación social también se utiliza, por ejemplo, para predecir fenómenos, que van desde medidas económicas hasta diversas cuestiones del comportamiento humano.

4.6.1 Modelo de Regresión Lineal

La regresión es lineal cuando la función de ajuste seleccionada es una recta, se habla también de recta de regresión.

La función de regresión a obtener es como se muestra en la ecuación

(4.4):

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon \quad (4.4)$$

Se llama modelo de regresión lineal múltiple con k número regresores x y una variable de respuesta y . Cuando existe una sola variable regresora, entonces se trata de un modelo de regresión lineal simple. El término ε , hace referencia al error aleatorio o residual; la diferencia entre el valor observado y_i y el valor ajustado \hat{y}_i se llama residual.

$$e_i = y_i - \hat{y}_i \quad i = 1, 2, \dots, n \quad (4.5)$$

Los parámetros $\beta_{j=0,1,\dots,k}$ se llaman coeficientes de regresión de las variables regresoras x_j , y representa el cambio esperado en la respuesta y por un cambio unitario en x_j , cuando todas las demás variables regresoras x_i ($i \neq j$) se mantienen constantes. El β_0 , representa la ordenada al origen.

Es importante para el análisis estimar los parámetros desconocidos en el modelo de regresión, pues de esta forma se puede adecuar el modelo a los datos.

4.6.2 Estimación de parámetros

Existen un grupo de técnicas especializadas en la estimación de parámetros. Dos métodos comúnmente empleados para para estimar los parámetros poblacionales de una muestra aleatoria son el método de estimación de máxima verosimilitud y el método de estimación de mínimos cuadrados (Mohaideen et al., 2015).

- **Mínimos Cuadrados**

Las estimaciones de los mínimos cuadrados (MC) se calculan mediante el ajuste de una línea de regresión a los puntos de un conjunto de datos que tiene la suma mínima de las desviaciones elevada al cuadrado (error de mínimos cuadrados).

Este método considera un grupo de garantías para validar los residuos, como: el coeficiente de correlación (r), el coeficiente de determinación (R^2), el mapa de residuos y el nivel de significancia o valor p . Su objetivo consiste en minimizar la suma de los residuos al cuadrado; a diferencia de la regresión cuantil, que se ve como minimizar una suma de errores absolutos ponderados con pesos asimétricos (Vicéns Otero et al., 2012).

De tal modo que la suma de cuadrados de los errores e_i , dada por la ecuación (4.6), se minimice.

$$S = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (4.6)$$

- **Método de la estimación de máxima verosimilitud (MLE)**

La función de verosimilitud señala la probabilidad de que una muestra observada dependa de los posibles valores de los parámetros. Estos parámetros son obtenidos al maximizar la función de verosimilitud y se

considera que son los que poseen mayor probabilidad de dar como resultado los datos observados. Este método se sugiere utilizar cuando existen muestras de gran tamaño, producto a su versatilidad, se puede aplicar a la mayoría de los modelos y a diferentes tipos de datos; además produce las estimaciones más precisas.

El MLE puede ser empleado incluso aun cuando existan datos censurados como parte de la muestra y además posee propiedades asintóticas con respecto a la distribución normal. Sus estimadores presentan varias propiedades, como los son:

- La consistencia
- Invariantes frente a transformaciones uno a uno, es decir, si θ y $g(\bar{\theta})$ es una función uno a uno de $\bar{\theta}$, entonces $g(\bar{\theta})$ es el estimador máximo verosímil de $g(\theta)$.
- Ser asintóticamente normales y eficientes, ya que entre todos los estimadores consistentes de un parámetro θ , los de máxima verosimilitud son los de varianza mínima.

Las propiedades esperadas de un estimador de parámetros son:

1. Ausencia de sesgo (ser insesgado).

Se dirá que $\hat{\theta}$ es un estimador insesgado de θ si la esperanza de $\hat{\theta}$ es θ .

Es decir, $E(\hat{\theta}) = \theta$.

2. Consistencia.

Se dice que un estimador es consistente si cumple:

$$\lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| > \varepsilon) = 0$$

Esta expresión indica que a medida que se incrementa el tamaño de la muestra, el valor del parámetro estimado se acerca al valor verdadero del parámetro.

3. Eficiencia.

Se emplea para comparar los estimadores. Si se tienen dos estimadores $\hat{\theta}_1$ y $\hat{\theta}_2$ de un mismo parámetro θ , se dirá que es más eficiente el que menor varianza tenga.

Mientras menor sea la estimación de la varianza, más pequeño será el tamaño de la muestra que se requiere para obtener el nivel de precisión deseado, y el estimador será más eficiente.

4. Suficiencia.

Se dice que $\hat{\theta}$ es un estimador suficiente del parámetro θ si dicho estimador basta por sí solo para estimar θ , usando toda la información disponible en el conjunto de datos.

La Verosimilitud proporciona un método general y versátil de estimación, permite censura, intervalos y datos truncados (Reyes Aguilar, 2006).

La censura es el fenómeno que ocurre cuando el valor de una observación sólo se conoce parcialmente, también puede ocurrir cuando hay observaciones fuera del rango de cierto instrumento de medida. Existen varios tipos de censura; se dice que es censura por intervalo cuando se sabe que una observación está entre dos valores extremos, pero no exactamente dónde; por la izquierda cuando una observación está por debajo de cierto valor, pero se ignora cuánto; y por la derecha si una observación está por encima de cierto valor, pero igualmente se ignora cuánto.

La contribución por diferentes tipos de censura para la función de distribución:

$$F(t) = \int_0^{t_i} f(t)dt \quad (4.7)$$

La verosimilitud en censura por intervalo es:

$$L_i(p) = \int_{t_{i-1}}^{t_i} f(t)dt = F(t_i) - F(t_i - 1) \quad (4.8)$$

La verosimilitud de censura por la izquierda es:

$$L_i(p) = \int_0^{t_i} f(t)dt = F(t_i) - F(0) = F(t_i) \quad (4.9)$$

La verosimilitud de censura por la derecha es:

$$L_i(p) = \int_0^{\infty} f(t)dt = F(\infty) - F(t_i) = 1 - F(t_i) \quad (4.10)$$

4.6.3 Validación del modelo

Validar un modelo es obtener un grado de confianza en el mismo para el fin al que se destina. No hay una “demostración” de la validez de un modelo, sino unos resultados positivos en un conjunto de test que brindan esa confianza. La validez de un modelo puede perderse por un solo test negativo.

- **Bondad de Ajuste (R^2)**

Una medida de ajuste que ha recibido gran aceptación en el contexto del análisis de regresión es el coeficiente de determinación R^2 . Se trata de una medida estandarizada que toma valores entre 0 y 1 (0 cuando las variables son independientes y 1 cuando entre ellas existe relación perfecta), es la proporción de la variación explicada por los regresores.

$$R^2 = 1 - \frac{SS_{res}}{SS_t} \quad (4.11)$$

La suma de cuadrados total (SS_t) es una medida de variabilidad de y sin considerar el efecto de las variables regresoras x . La suma de cuadrados de los residuales (SS_{res}) es una medida de la variabilidad de y después de haber considerado x . Los valores de R^2 cercanos a 1

implican que la mayor parte de la variabilidad y está explicada por el modelo de regresión.

- **Validación Cruzada**

La validación cruzada o cross-validation es una técnica utilizada para evaluar los resultados de un análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y prueba. Consiste en repetir y calcular la media aritmética obtenida de las medidas de evaluación sobre diferentes particiones. Se utiliza en entornos donde el objetivo principal es la predicción y se quiere estimar la precisión de un modelo que se llevará a cabo a la práctica. Es una técnica muy utilizada en proyectos de inteligencia artificial para validar modelos generados. Este método consiste en dejar una observación fuera del modelo y calcular el error o diferencia entre el valor observado y el valor predicho (Kohavi R, 1995).

La validación cruzada dejando uno fuera o *Leave-one-out cross-validation (LOOCV)* implica separar los datos de forma que para cada iteración se tenga una sola muestra para los datos de prueba y todo el resto conformando los datos de entrenamiento. La evaluación está dada por el error. A nivel computacional es muy costoso, puesto que se tienen

que realizar un elevado número de iteraciones, tantas como N muestras se tengan y para cada una analizar los datos tanto de entrenamiento como de prueba (Elkan, 2011).

En la validación cruzada de K iteraciones o K -fold cross-validation los datos de muestra se dividen en K subconjuntos. Uno de los subconjuntos se utiliza como datos de prueba y el resto ($K-1$) como datos de entrenamiento. El proceso de validación cruzada es repetido durante k iteraciones, con cada uno de los posibles subconjuntos de datos de prueba. Finalmente se realiza la media aritmética de los resultados de cada iteración para obtener un resultado único.

Esto significa que, el ajuste del modelo es realizado N (K) veces con la muestra de entrenamiento, omitiendo solo un valor de la muestra en cada ajuste. Para realizar la estimación del error, se realiza la predicción del punto excluido en el ajuste para cada caso.

Geisser (1993) plantea que la evaluación del error que se obtiene mediante este método es buena, sin embargo, tiene el inconveniente de ser costoso computacionalmente, por otra parte; tiene la ventaja de ser la mejor forma de evaluar un modelo de predicción. Para evaluar la predicción, puede ser usado el estadístico de predicción (PRESS) o la raíz del cuadrado medio del error (RMSPE).

Montgomery et.al., (2006) definieron los residuales PRESS como $e_i = y_i - \hat{y}_i$, siendo y_i el valor predicho de la i -ésima respuesta observada, basado en un ajuste de modelo con los $N - 1$ puntos de muestra. Se observó que los residuales PRESS grandes, tienen una utilidad potencial para identificar observaciones donde el modelo no se ajusta bien a los datos, o para observaciones en las que es probable que el modelo produzca malas predicciones futuras.

Allen (1971) ha sugerido el uso de la Suma de Cuadrados de Error de Predicción (es decir, la estadística PRESS), definida como la suma de los residuales PRESS al cuadrado, como medida de la calidad del modelo. La estadística PRESS es:

$$PRESS = \sum_{i=1}^N [y_i - \hat{y}_{(i)}]^2 \quad (4.12)$$

Con la estadística PRESS se puede calcular un estadístico parecido a la R^2 para predicción.

$$R^2_{PRESS} = 1 - \frac{PRESS}{SS_t} = 1 - \frac{\sum_{i=1}^N [y_i - \hat{y}_{(i)}]^2}{\sum_{i=1}^N [y_i - \bar{y}]^2} \quad (4.13)$$

Este indicador brinda una medida del grado de fiabilidad en la predicción de nuevas observaciones.

4.7 Regresión Robusta

Cuando las observaciones y en el Modelo de Regresión Lineal tienen una distribución (considerablemente) no normal y/o existen valores atípicos que afectan al modelo de regresión, no es conveniente utilizar el método de mínimos cuadrados ordinarios (MC) como procedimiento de estimación de parámetros. Un caso de interés práctico es aquél en el que las observaciones tienen una distribución que tiene colas más largas o gruesas que la distribución normal. Esas distribuciones tienden a generar valores atípicos, que pueden tener una gran influencia sobre el método de los mínimos cuadrados, al orientar demasiado la ecuación de regresión en su dirección.

$$y_i = X\beta + \varepsilon \quad (4.14)$$

Una vía para solucionar esta situación sería descartar los valores atípicos, lo cual resultaría ideal desde el punto de vista estadístico pues se obtiene una recta que se adecua a los demás valores. En este caso, no es la decisión más acertada pues se estaría perdiendo una parte de la información, lo que conlleva a que se puedan cometer errores; además en algunas situaciones donde están implicadas un gran número de

variables explicativas y la muestra es mayor, resulta complicado identificar si el modelo de regresión se ha distorsionado por las observaciones atípicas (Montgomery et al., 2006).

Cuando existen datos atípicos los métodos de regresión robustos constituyen técnicas potenciales para usar, ya que al contar con la robustez de sus estimadores robustos es posible ponderar los valores que se posicionan más distantes de los centrales en una serie de datos.

Un procedimiento de regresión robusto es aquel que amortigua el efecto de las observaciones que serían muy influyentes si se usaran los mínimos cuadrados, lo que indica que un procedimiento robusto tiende a dejar grandes los residuales asociados con valores atípicos, facilitando así la identificación de puntos influyentes. Además de la insensibilidad a los valores atípicos, un procedimiento de estimación robusta debería producir, en esencia, los mismos resultados obtenidos por el método MC cuando la distribución básica es normal, y cuando no hay valores atípicos (Carmonab et al., 2015).

La idea básica de la regresión robusta es utilizar un estimador lo suficientemente fuerte, que, a través de una función de ponderación introducida, pueda reducir (e incluso eliminar) el efecto de los residuos altos. Por tanto, se definen los pesos de forma que tomen valores pequeños en los residuos “grandes”.

4.8 Regresión Cuantil

La regresión cuantil puede considerarse un caso particular de la regresión robusta. Es una generalización del concepto de cuantil. Fue presentada por Koenker y Bassett (1978) como una extensión de la estimación de los mínimos cuadrados clásicos de los modelos medios condicionales a funciones cuantiles condicionales.

Buchinsky (1998) expone dos ventajas más del modelo de regresión cuantil. Primero que posee estimadores más eficientes que los estimadores mínimos cuadrados, ya que estos últimos tienen como condición restrictiva que requieren el cumplimiento de la distribución idéntica e independiente para los errores. Si la distribución de errores no es normal, los estimadores de la regresión cuantil resultan más eficientes en comparación con los de mínimos cuadrados. Por otra parte, estos estimadores tienen la característica de robustez pues la función objetivo empleada para estimar los parámetros en regresión cuantil se basa en una suma ponderada de desviaciones absolutas. La regresión cuantil generaliza el concepto de cuantil de una variable aleatoria a una función cuantil, dados los valores observados de un conjunto de covariables.

La regresión cuantil establece los modelos de cuantiles condicionales como funciones de predicción, constituyendo una extensión natural del modelo de regresión lineal. Mientras que el modelo de regresión lineal especifica el cambio en la media condicional de la variable dependiente asociada con un cambio en las covariables, el modelo de regresión cuantil especifica los cambios en el cuantil condicional. Dado que cualquier cuantil se puede utilizar, es posible modelar cualquier posición predeterminada de la distribución. Por lo tanto, los investigadores pueden elegir posiciones que se adaptan a sus investigaciones específicas.

4.8.1 Cuantil

Así como la regresión por MC se encuentra vinculada con la media como se veía en la introducción, la regresión cuantil, como su propio nombre deja entrever, se basa en el concepto de cuantil.

Los cuantiles son puntos tomados a intervalos de la función de distribución de una variable aleatoria, dividiendo la misma en partes. La Función de Distribución Acumulada (CDF) es aquella F_Y que da para cada valor de y la proporción de la población para la cual $Y \leq y$.

$$P(Y \leq y) \geq F_Y(y) \quad (4.15)$$

$$P(Y \geq y) \geq 1 - F_Y(y) \quad (4.16)$$

Es decir, la probabilidad de observar un valor por debajo (o igual que) y debe ser (al menos) F_Y , mientras que la probabilidad de observar un valor por encima (o igual que) y debe ser (al menos) $1 - F_Y$ (Fahrmeir et.al., 2013).

El p -ésimo cuantil de la distribución denotado como $Q^{(p)}$, es el valor de la inversa de la distribución de probabilidad acumulada (CDF) de p ; esto es, un valor de y tal que:

$$F(y) = p \tag{4.17}$$

Es decir, se tendrá que el cuantil $Q^{(p)}$ de la muestra, con $0 < p < 1$, será aquel $Q^{(p)}$ valor que deje una proporción p de observaciones por debajo de $Q^{(p)}$ y una proporción $(1 - p)$ por encima. En el caso de la mediana (Me) $p = 0.5$, quedarán un 50% de los datos por debajo de $Q^{(p)} = Me$ y un 50% de los datos por encima. Si se utiliza el primer cuartil ($p = 0.25$) sería un 25% de los valores de Y los que quedarían por debajo de $Q^{(p)} (p = Q_1)$ y un 75% por encima, y de forma similar e inversa con el tercer cuartil. Los cuartiles dividen la muestra en 4 partes, pero de igual manera se puede dividir la muestra en 10 partes con los deciles, $p = 0,1; 0,2; \dots, 0,9$ o cualquier otra proporción. Estas

partes es lo que se denominan cuantiles, de los que la mediana, los cuartiles o deciles son casos particulares (Sánchez et al., 2012).

Un cuantil se puede utilizar para calcular la proporción para cualquier intervalo de valores de y , como refleja la ecuación (4.18).

$$P[a < Y \leq b] = F_Y(b) - F_Y(a) \quad (4.18)$$

Para una muestra grande, representada con una distribución de función cuantil, $Q^{(p)}$ y la función de densidad de probabilidad $F = f'$; la distribución de $\hat{Q}^{(p)}$ es aproximadamente normal con media $Q^{(p)}$ y varianza $\frac{p(1-p)}{n} \cdot \frac{1}{f(Q^{(p)})^2}$.

Para estimar la variabilidad de la muestra cuantil, se hace necesario el uso de la aproximación de la varianza. Un enfoque estándar para la estimación donde la pendiente de la recta tangente a la función en el punto p ; es la derivada de la función cuantil con respecto a p ; o de manera equivalente; la inversa de la función de densidad:

$$\frac{d}{dp} Q^{(p)} = 1/f(Q^{(p)}) \quad (4.19)$$

Este término puede ser aproximado por la diferencia del cociente:

$$\frac{1}{2h} (\hat{Q}^{(p+h)} - \hat{Q}^{(p-h)}) \quad (4.20)$$

Que es la pendiente de la línea que pasa por los puntos: $(p - h, \hat{Q}^{(p-h)})$ y $(p + h, \hat{Q}^{(p+h)})$, para un valor pequeño de h , como se muestra en la figura 4.3.

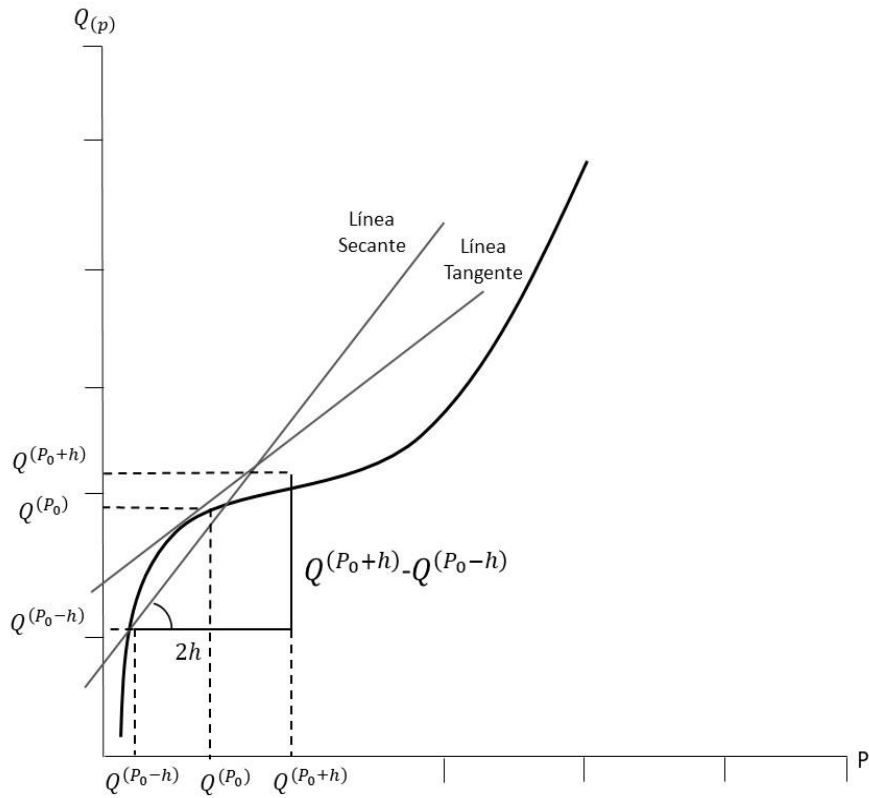


Figura 4.3. Ilustración de como estimar la pendiente de una función cuantil.

4.8.2 Modelo de Regresión Cuantil

El Modelo de Regresión Cuantil (QRM), estima el potencial efecto diferencial de una covariable en varios cuantiles en la distribución condicional. Con la mediana y los cuantiles fuera de la mediana, estas líneas de regresión ajustada, capturan el cambio de ubicación (la línea

de la mediana), así como la escala y los cambios de forma más complejos (las líneas para cuantiles fuera de la mediana). De esta manera, el QRM estima el efecto diferencial de una covariable en la distribución completa y tiene capacidad para heterocedasticidad. Siguiendo Koenker y Bassett (1978), el QRM correspondiente al LRM (método de regresión lineal) en la ecuación (4.21) se puede expresar como:

$$\hat{y}_i = \beta_0^{(p)} + \beta_1^{(p)} x_i + \varepsilon_i^{(p)} \quad (4.21)$$

Donde $0 < p < 1$, indicando la proporción de la población que tiene valores por debajo del cuantil p .

Resumiendo, la regresión cuantil ofrece la posibilidad de crear distintas rectas de regresión para distintos cuantiles de la variable dependiente a través de un método de estimación, que se representa como (Fahrmeir et.al.,2013):

$$y_i = x'_i \beta_p + \varepsilon_{ip} \quad i = 1, \dots, n, \quad (4.22)$$

Donde:

y_i : es la variable dependiente.

x'_i : representa a la matriz de variables independientes.

β_p : es el parámetro a estimar correspondiente al cuantil p .

ε_{ip} : es la perturbación aleatoria correspondiente al cuantil p .

El QRM se puede formular de manera equivalente al modelo de regresión lineal (LRM). Solo que en este caso, se especifica el p –ésimo cuantil condicional dado x_i , como se muestra en la ecuación (4.26). A diferencia del LRM que tiene solo una medida condicional, el QRM puede tener numerosos cuantiles condicionales.

Dado que la expectativa de que el error ε_i tiene valor esperado cero en los modelos de regresión lineal, también en el QRM se puede formular de manera equivalente con una declaración acerca de los términos de error.

$$Q^{(p)}(y_i | x_i) = \beta_0^{(p)} + \beta_1^{(p)} x_i + Q^p(\varepsilon_i) = \beta_0^{(p)} + \beta_1^{(p)} x_i \quad (4.23)$$

Es importante señalar que para diferentes valores del p –ésimo cuantil de interés, los términos del error ε_i para una i fija, están relacionados. También supone que cada uno de los errores $\varepsilon_i^{(p)}$ para $i = 1, \dots, n$ son independientes e idénticamente distribuidos.

4.8.3 Estimación

Una diferencia significativa entre el estimador de la regresión cuantil y el de regresión lineal por mínimos cuadrados, es que la distancia entre

los puntos de la línea ajustada se mide usando una suma ponderada de las distancias verticales.

En el caso de regresión mediana, se elige minimizar la suma de residuos absolutos. En otras palabras, encontrar los coeficientes que minimizan la suma de los residuales absolutos (la distancia absoluta desde un valor observado, hasta su valor ajustado).

El estimador resuelve las β s minimizando la siguiente ecuación:

$$\sum_i |y_i - \beta_0 - \beta_1 x_i| \quad (4.24)$$

De manera similar en la regresión cuantil, se define los p – *ésimos* estimadores de regresión cuantil $\hat{\beta}_0^{(p)}$ y $\hat{\beta}_1^{(p)}$ como los valores que minimizan la suma ponderada de las distancias entre valores ajustados $\hat{y}_i = \hat{\beta}_0^{(p)} + \hat{\beta}_1^{(p)} x_i$ y y_i , donde el peso es $1 - p$ para los puntos por debajo de la línea y p para los puntos sobre la línea. Cada elección de esta proporción de p , por ejemplo, $p = .10, .25, .50$, da lugar a una diferente función ajustada de cuantiles condicionales.

En otras palabras se busca minimizar una suma ponderada de los residuales $y_i - \hat{y}_i$.

$$\sum_{i=1}^n d_p(y_i, \hat{y}_i) = p \sum_{y_i \geq \beta_0^p + \beta_1^p x_i} |y_i - \beta_0^p - \beta_1^p x_i| + (1-p) \sum_{y_i < \beta_0^p + \beta_1^p x_i} |y_i - \beta_0^p - \beta_1^p x_i| \quad (4.25)$$

Donde : d_p , es la distancia desde y hasta q .

$$d_p(y,q) = \begin{cases} (1-p)|y-q| & y < q \\ p(y-q) & y \geq q \end{cases}$$

Por ejemplo, cuando estiman los coeficientes para $p = 0.10$ línea de regresión cuantil, las observaciones debajo de la línea tienen un peso de 0.90 y los que están sobre la línea reciben un peso menor de 0.10. Como resultado, 90% de los puntos de datos (x_i, y_i) se encuentran por encima de la línea ajustada que conduce a residuos positivos, y el 10% está debajo de la línea y , por lo tanto, tiene residuos negativos. Por el contrario, para estimar los coeficientes para la regresión cuantil $p = 0.90$, puntos debajo de la línea se les da un peso de 0.10, y el resto tiene un peso de 0.90; como resultado. El 90% de las observaciones tienen residuos negativos y el 10% restante residuos positivos.

4.8.4 Inferencia en la regresión Cuantil

Dentro del marco de la regresión cuantil dada la literatura, se muestran dos variantes empleadas para realizar estimaciones, el método denominado Bootstrap y la estimación no paramétrica de funciones de

densidad, mediante el método Kernel, la cual permite construir una función de densidad. En este caso se estará abordando el método de Bootstrap.

- **El método Bootstrap para la QRM**

Se introduce el procedimiento de Bootstrap para realizar las inferencias acerca de los coeficientes de QRM.

El Bootstrap es una técnica de remuestreo que constituye una alternativa para estimar la distribución de un parámetro estimado, a partir de una muestra de tamaño n , permitiendo aproximar la distribución empírica de los parámetros estimados a una distribución de muestreo deseada.

Para una muestra de tamaño n se realiza un remuestreo, reemplazando el actual conjunto de datos por un número de muestras de reemplazo, llamado M . Generalmente entre 50 y 200 para estimar la desviación estándar; y entre 2000 y 5000 para un intervalo de confianza. Cada remuestreo tendrá la misma cantidad que la muestra original, y puede incluir los datos originales más de una vez.

Un enfoque para esto es obtener una aproximación de la desviación estándar (S_{boot}), según la ecuación (4.26).

$$\sqrt{\frac{p(1-p)}{nf(Q^{(p)})^2}} \quad (4.26)$$

Donde se puede estimar el término, $1/f(\hat{Q}^{(p)})$ utilizando $(\hat{Q}^{(p+h)} - \hat{Q}^{(p-h)})/(2h)$, para una elección apropiada de la variable h . Las estimaciones de bootstrap también se pueden usar para formar un intervalo de confianza apropiado para la población del p –ésimo percentil. Una posibilidad es hacer uso de la estimación original $\hat{Q}^{(p)}$, su error estándar estimado S_{boot} y la aproximación normal dando un intervalo de confianza de $100(1 - \alpha)\%$ de la forma $\hat{Q}^{(p)} \pm Z_{\alpha/2}S_{boot}$.

También puede ser empleado para producir intervalos de confianza para un parámetro de regresión cuantil individual $\hat{\beta}_i^{(p)}$, de la forma $\hat{\beta}_i^{(p)} \pm Z_{\alpha/2}S_{boot}$.

Se puede estimar la covarianza entre todos los posibles coeficientes de regresión cuantil de cada uno de los modelos, por ejemplo, cuando el modelo que se está ajustando contiene un parámetro de intercepto $\hat{\beta}_0^{(p)}$ y coeficientes correspondientes a dos covariables $\hat{\beta}_1^{(p)}$ y $\hat{\beta}_2^{(p)}$, con los cuales se forma una matriz de covarianza. Esta matriz proporciona no

solo la varianza para el coeficiente de cada covariable en cada cuantil, sino también la covarianza de las estimaciones en diferentes cuantiles para la misma covariable.

Con la varianza y la covarianza estimadas, se puede realizar prueba de hipótesis en la equivalencia para un par de coeficientes $\beta_i^{(p)}$ y $\beta_i^{(q)}$ correspondiente a la misma covariable, pero a través de distintos cuantiles p y q usando el estadístico de Wald:

$$W = \frac{(\hat{\beta}_j^{(p)} - \hat{\beta}_j^{(q)})^2}{\hat{\sigma}_{\hat{\beta}_j^{(p)} - \hat{\beta}_j^{(q)}}^2} \quad (4.27)$$

El término del denominador es la varianza estimada de la diferencia $\hat{\beta}_j^{(p)} - \hat{\beta}_j^{(q)}$, que se obtiene utilizando la igualdad (4.28), sustituyendo las varianzas y covarianzas estimadas en el lado derecho:

$$Var(\hat{\beta}_j^{(p)} - \hat{\beta}_j^{(q)}) = Var(\hat{\beta}_j^{(p)}) + Var(\hat{\beta}_j^{(q)}) - 2Cov(\hat{\beta}_j^{(p)}, \hat{\beta}_j^{(q)}) \quad (4.28)$$

Bajo la hipótesis nula, el estadístico de Wald tiene una distribución aproximada Ji-cuadrada X^2 con un grado de libertad. En términos más generales, se puede probar la igualdad de múltiples coeficientes a través de cuantiles. Por ejemplo, suponiendo que se tienen dos covariables

además del término de intercepción en los modelos, es posible que se desee probar si el condicional las funciones cuantiles p –ésimo y q –ésimo son cambios entre sí; es decir:

$$H_0: \beta_i^{(p)} = \beta_i^{(q)} \text{ y } \beta_j^{(p)} = \beta_j^{(q)}$$

$$H_a: \beta_i^{(p)} \neq \beta_i^{(q)} \text{ y } \beta_j^{(p)} \neq \beta_j^{(q)}$$

El estadístico de Wald para realizar esta prueba se puede describir de la siguiente manera. Primero, se emplean las covarianzas estimadas para obtener una matriz de covarianza estimada,

$$\Sigma_{\hat{\beta}^{(p)} - \hat{\beta}^{(q)}} = \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} \\ \hat{\sigma}_{21} & \hat{\sigma}_{22} \end{bmatrix} \quad (4.29)$$

donde las entradas se obtienen sustituyendo varianzas y covarianzas estimadas. Los términos de la covarianza se obtienen mediante la ecuación (4.30):

$$\hat{\sigma}_{12} = \hat{\sigma}_{21} \text{Cov}(\hat{\beta}_1^{(p)}, \hat{\beta}_2^{(p)}) \text{Cov}(\hat{\beta}_1^{(q)}, \hat{\beta}_2^{(q)}) \text{Cov}(\hat{\beta}_1^{(p)}, \hat{\beta}_2^{(q)}) \text{Cov}(\hat{\beta}_1^{(q)}, \hat{\beta}_2^{(p)}) \quad (4.30)$$

Luego se puede calcular la estadística de prueba como:

$$W = \begin{bmatrix} \hat{\beta}_1^{(p)} - \hat{\beta}_1^{(q)} \\ \hat{\beta}_2^{(p)} - \hat{\beta}_2^{(q)} \end{bmatrix}^t \sum_{\hat{\beta}^{(p)} - \hat{\beta}^{(q)}}^{-1} \begin{bmatrix} \hat{\beta}_1^{(p)} - \hat{\beta}_1^{(q)} \\ \hat{\beta}_2^{(p)} - \hat{\beta}_2^{(q)} \end{bmatrix} \quad (4.31)$$

Presentando una distribución aproximadamente Ji-cuadrada X^2 con dos grados de libertad, permite a los investigadores probar la equivalencia de coeficientes a través de cuantiles.

- **Error estándar e intervalos de confianza**

Se puede obtener el error estándar y el intervalo de confianza para cualquier estimación y combinaciones de estimaciones. Permite hacer inferencias para los coeficientes $\beta^{(p)}$ en el QRM escritos en la forma $Q^{(p)}(y_i, x^{(i)}) = \sum_{j=1}^k \beta_j^{(p)} x_j^{(i)}$. Una forma equivalente de esto indica que $y_i = \sum_{j=1}^k \beta_j^{(p)} x_j^{(i)} + \varepsilon_i^{(p)}$, donde el $\varepsilon_i^{(p)}$ tiene una distribución común cuyo p – ésimo cuantil es cero.

La inferencia para un coeficiente $\beta_j^{(p)}$ lo hará estar en la forma de un intervalo de confianza o una prueba de hipótesis basada en alguna medida del error estándar $S_{\hat{\beta}_j^{(p)}}$ de $\hat{\beta}_j^{(p)}$, como en la configuración LRM. El error estándar tiene la propiedad de que asintóticamente, la

cantidad $\left(\hat{\beta}_j^{(p)} - \beta_j^{(p)} \right) / S_{\hat{\beta}_j^{(p)}}$ tiene una distribución normal estándar.

Los errores estándar para el QRM son considerablemente más simples y fáciles de describir. En este caso, la matriz asintótica de covarianza para $\hat{\beta}_j^{(p)}$ toma la forma:

$$\sum_{\hat{\beta}^{(p)}} = \frac{p(1-p)}{n} \cdot \frac{1}{f_{\varepsilon^{(p)}}(0)^2} (X^t X)^{-1} \quad (4.32)$$

El término $f_{\varepsilon^{(p)}}(0)^2$ constituye la densidad de probabilidad del término de error $\varepsilon^{(p)}$ evaluado en el p –ésimo cuantil de la distribución del error. Como en la LRM, la matriz de covarianza es un múltiplo escalar de la matriz $(X^t X)^{-1}$. Sin embargo, en QRM, el multiplicador $\frac{p(1-p)}{n} \cdot \frac{1}{f_{\varepsilon^{(p)}}(0)^2}$ es la varianza asintótica de un cuantil basado en una muestra (univariada) $\varepsilon_1^{(p)}, \dots, \varepsilon_n^{(p)}$.

El término de densidad que aparece en esta expresión es desconocido y debe ser estimado, a partir de $\frac{1}{f_{\varepsilon^{(p)}}} = \frac{d}{dp} Q^{(p)}(\varepsilon^{(p)})$ usando un cociente de diferencia, como se muestra en la ecuación (4.33), donde los cuantiles de muestra $\hat{Q}^{(p)}(p \pm h)$ se basan en los residuales de la ecuación (4.34) para el modelo de QRM ajustado.

$$\frac{1}{2h} (\hat{Q}^{(p)}(p+h) - \hat{Q}^{(p)}(p-h)) \quad (4.33)$$

$$\hat{\varepsilon}_i^{(p)} = y_i \sum_{j=1}^k \hat{\beta}_j^{(p)} x_j^{(i)}, \quad i = 1, \dots, n \quad (4.34)$$

La elección de h para usar es delicada; Koenker (2005) describe algunos enfoques para su elección.

Con un estimador consistente de la varianza se puede construir un intervalo de confianza para los parámetros estimados, de la forma:

$$IC = \left[\hat{\beta}_j^{(p)} - z_{1-\alpha/2} \sqrt{\hat{V}_{ar}} ; \hat{\beta}_j^{(p)} + z_{1-\alpha/2} \sqrt{\hat{V}_{ar}} \right] \quad (4.35)$$

- **Bondad de Ajuste de la QRM**

Una vez obtenida la matriz de varianzas y covarianzas, ya se pueden hacer contrastes sobre la nulidad de los parámetros de la regresión, ahora bien, como medida global de ajuste de la misma. Para la regresión cuantil, se desarrolla un análogo del estadístico R^2 utilizado en la regresión lineal.

Dado que los ajustes del modelo de regresión lineal se basan en el criterio de mínimos cuadrados, y los modelos de regresión por cuantiles se basan en minimizar la suma de distancias ponderadas, se requiere

medir la bondad del ajuste de una manera que sea consistente con este criterio.

Koenker y Machado (1999) sugieren medir la bondad del ajuste comparando la suma de las distancias ponderadas para el modelo de interés, con la suma en la que solo aparece el parámetro de intersección, como se muestra la ecuación (4.35). Dado $V^1(p)$ como la suma de las distancias ponderadas para el p –ésimo modelo de la regresión cuantil completa, y $V^0(p)$ como la suma de la distancia ponderada para el modelo que incluye solo un término constante.

$$R(p) = 1 - \left(\frac{V^1(p)}{V^0(p)} \right) \quad (4.36)$$

Donde:

$$V^1(p) = \sum_{i=1}^n d_p(y_i, \hat{y}_i) = p \sum_{y_i \geq \beta_0^p + \beta_1^p x_i} |y_i - \beta_0^p - \beta_1^p x_i| + (1-p) \sum_{y_i < \beta_0^p + \beta_1^p x_i} |y_i - \beta_0^p - \beta_1^p x_i|$$

$$V^0(p) = \sum_{i=1}^n d_p(y_i, \hat{Q}^{(p)}) = \sum_{y_i \geq \bar{y}} p |y_i - \hat{Q}^{(p)}| + \sum_{y_i < \bar{y}} (1-p) |y_i - \hat{Q}^{(p)}|$$

Planteado de otra manera, el numerador representa la suma de residuos en valor absoluto y el denominador la suma de las desviaciones de cada

valor real de la variable endógena al cuantil p muestral (Sánchez & Vicéns, 2012).

$$Pseudo - R^2 = 1 - \frac{\sum_{i=1}^n |y_i - \hat{y}_1|}{\sum_{i=1}^n |y_i - y_{\theta}|} \quad (4.36)$$

Esta medida de ajuste, denominada pseudo- R^2 , para distinguirla de la R^2 de la regresión lineal, tiene un valor que se encuentra dentro del rango de 0 a 1, mientras más se aproxime a 1, indica un mejor ajuste del modelo. La ecuación anterior representa una medida local de la bondad del ajuste del modelo en p . La evaluación global de un QRM para toda la distribución requiere un examen del $R(p)$ colectivamente.

4.9 Optimización multiobjetivo

Siempre encontrar la solución útil de un problema cuyo espacio de búsqueda es muy amplio, implica la existencia de limitaciones de tiempo, además de que no exista un método analítico mediante el cual se pueda llegar a una solución exacta.

Los problemas reales usualmente requieren la búsqueda de soluciones que satisfagan en forma simultánea múltiples criterios de desempeño u objetivos que pueden ser contradictorios, es decir, se desconoce la manera óptima de combinar los diferentes objetivos o que sea

inadecuado, cuando no es imposible hacerlo, a lo cual se le conoce como un Problema de Optimización Multiobjetivo (MOP en inglés de Multiobjective Optimization Problem) (Zitzler et al., 2000).

Un problema de optimización multiobjetivo (MOP) es aquel que incluye un conjunto de n variables de decisión, un conjunto de k funciones objetivo, y un conjunto de m restricciones de desigualdad y p restricciones de igualdad, en donde las funciones objetivo y las restricciones son funciones de las n variables de decisión (Díaz et al., 2002).

En general un problema de optimización multiobjetivo se formula como (Peñuela et al., 2007):

$$\min/\max \quad f_m(x) \quad m = 1, 2, \dots, M \quad (4.37)$$

Sujeto a:

$$g_i(x) \geq 0 \quad i = 1, 2, \dots, I$$

$$h_k = 0 \quad k = 1, 2, \dots, K$$

$$x_j^l \leq x_j \leq x_j^u \quad j = 1, 2, \dots, J$$

En este caso se tiene un grupo de funciones para ser optimizadas de forma simultánea, lo que marca la diferencia con los modelos mono-objetivos. El objetivo es establecer un vector de variables de estado

$x = (x_1, x_2, \dots, x_j)$ que cumpla con en el conjunto restricciones y donde las funciones objetivos resultantes sean optimizadas.

Todas las combinaciones posibles en el valor de las variables, que representan el espacio de solución, genera un segundo espacio vectorial conocido como “Espacio Objetivo”.

Cuando se trata con problemas de optimización mono-objetivo una alternativa de solución se considera mejor que otra si produce una solución objetivo de menor valor para el caso de minimización, o de mayor valor para el caso de maximización. Sin embargo, en los problemas multiobjetivo este criterio debe ser revaluado, ya que se considera al mismo tiempo funciones de minimizar y funciones de maximizar (Peñuela et al., 2007).

Los algoritmos evolutivos utilizan mecanismos inspirados en la evolución biológica, tales como la selección, reproducción y la mutación. Las soluciones se modelan como individuos de una población. Una función de aptitud determina el entorno donde se encuentran las soluciones. Dentro de los algoritmos evolutivos están los algoritmos genéticos (AG), que son una técnica de búsqueda iterativa, buscando derivar estrategias de optimización basadas en la generación

de poblaciones de individuos aptos mediante la reproducción de los padres (Holland, 1988).

4.9.1 Algoritmo NSGA-II

En el año 2000 fue propuesto el algoritmo NSGA-II (Elitist Non-Dominated Sorting Genetic Algorithm). En este algoritmo, la población descendiente Q_t (tamaño N) es creada en primera instancia usando la población de padres P_t (tamaño N).

Posteriormente, se combinan ambas poblaciones y se obtiene R_t de tamaño $2N$, la cual es clasificada en diferentes frentes de Pareto a través de un ordenamiento no dominado. Esto permite una verificación global de dominancia entre la población de padres y descendientes.

Cuando el proceso de ordenamiento no dominado ya esté finalizado, se genera la nueva población a partir de las configuraciones de los frentes no dominados. Esta comienza a ser construida con el mejor frente no dominado (F_1), continúa con las soluciones del segundo frente (F_2), luego del tercero (F_3) y así sucesivamente. Sin embargo, no es posible acomodar en la nueva población todas configuraciones de los frentes correspondientes a R_t , ya que esta población es de tamaño $2N$, y solo hay N configuraciones que conforman la población descendiente.

Aquellos frentes que no pueden ser acomodados desaparecen (Deb et al., 2000).

Resulta útil utilizar una maniobra para elegir las configuraciones situadas en un área alejada de otras soluciones o poco poblada para completar las posiciones restantes de la población descendiente a cambio de elegir configuraciones aleatoriamente, ya que, cuando se está considerando el último frente, las soluciones que forman parte de él, pueden exceder las restantes por acomodar en la población descendiente, dicha situación se ilustra en la figura 4.4.

Es necesario que las configuraciones no rechazadas sean escogidas mediante una metodología que garantice la diversidad y que además asegure su calidad, puesto que, con el avance del proceso, algunas configuraciones pasan a formar parte del primer frente, el cual muchas veces llega a tener más de N individuos.

Con el algoritmo NSGA-II se obtiene en un inicio una población (aleatoria o mediante una técnica de inicialización) de padres P_0 . Esta población es ordenada según corresponda a los niveles de no dominancia (ordenamiento de los frentes de Pareto, F_1, F_2, \dots). Posteriormente se le asigna una función “fitness” a cada una de las

soluciones según su nivel de no dominancia (1 es el mejor nivel) y se supone que durante el proceso de debe decrecer dicha función.

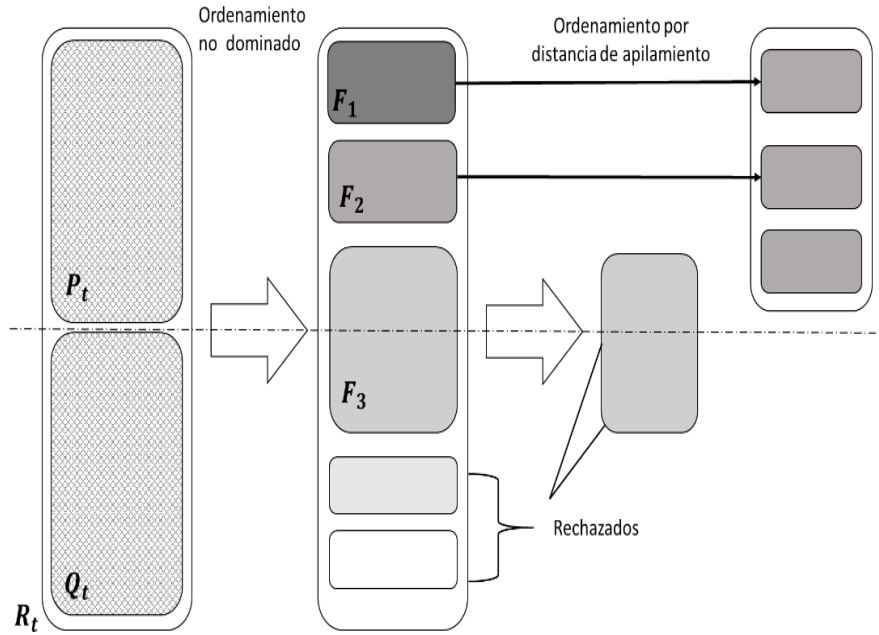


Figura 4.4. Diagrama esquemático del mecanismo de promoción de individuos del NSGA-II (Duardo Toste, 2015).

La selección por torneo (empleando un operador de torneo para apilamiento, descrito posteriormente), el cruzamiento y la mutación son utilizados para crear la población de descendientes Q_0 de tamaño N . Los principales pasos del algoritmo NSGA-II son descritos a continuación (Duardo Toste, 2015):

1. Combinar las poblaciones de padres y descendientes para crear $R_t = P_t \cup Q_t$. Realizar el ordenamiento no dominado a R_t e identificar los frentes $F_i, i = 1, 2, \dots, etc.$

2. Hacer $P_{i+t} = \emptyset$, e $i = 1$. Mientras $|P_{i+1}| + |F_i| \leq N$ hacer $|P_{i+1}| + |P_{+1}| \cup |F|$ e identificar los frentes $i = i + 1$.

3. Realizar ordenamiento por apilamiento ($F'_i < C$) e incluir en P_i , las $N - |P_{i+1}|$ soluciones más esparcidas, usando los valores de distancia de apilamiento asociadas al frente F_i .

Este operador ($F'_i < C$) compara dos soluciones y elige un ganador del torneo. Se considera que una solución tiene un rango de no dominancia asociado (r_i) y una distancia de apilamiento (d_i). La d_i de una solución i , es la medida del espacio de búsqueda alrededor de i que no está ocupado por otra solución en la población.

Por medio de estos dos atributos, es posible hallar la mejor configuración a través del operador de selección por torneo para apilamiento, siendo esta la que tenga mejor rango de no-dominancia. En el caso de estar situadas en el mismo frente de Pareto la ganadora del torneo será la que posea la mejor distancia de apilamiento, es decir la más alta. Esta distancia (d_j para cada solución j), según un índice I , es determinada algorítmicamente, haciendo uso de la ecuación (4.38):

$$d_j^m = d_j^m + \frac{f_m^{(I_{j+1}^m)} - f_m^{(I_{j-1}^m)}}{f_m^{max} - f_m^{min}} \quad (4.38)$$

En dicho caso , f_m^{max} , f_m^{min} representan el máximo y mínimo valor respectivamente de la función objetivo m . Mientras, $f_m^{(I_{j+1}^m)}$, $f_m^{(I_{j-1}^m)}$ constituyen las soluciones vecinas a la configuración j para cada una de las funciones objetivo m .

Las distancias consideran todas las funciones objetivo y se asigna el valor de infinito a las soluciones extremas del frente de Pareto considerado. La distancia resultante se obtiene a partir de la adición de las distancias en cada una de las direcciones de las funciones objetivo del problema, debido a que son las que tienen el mejor valor de una de las funciones objetivo del frente. De esta forma el algoritmo impulsa las configuraciones que ocupen los mejores frentes hacia el próximo ciclo generacional, utilizando los operadores genéticos básicos.

4. Crear la población descendiente Q_{i+1} a partir de P_{i+1} usando selección por torneo para apilamiento, cruzamiento y mutación.

Así, de este modo, el algoritmo impulsa hacia el siguiente ciclo generacional las configuraciones más diversas que llenen los mejores frentes, a través de los operadores genéticos básicos.

CAPÍTULO 5

5. Metodología

Considerando el problema planteado en el capítulo 2 y la revisión de la literatura referente a los modelos de regresión cuantil y el proceso de PTA, el presente capítulo se enfoca en la metodología propuesta para construir un modelo que sea capaz de representar el desgaste de la herramienta tratada superficialmente mediante el proceso PTA.

5.1 Metodología Propuesta

La figura 5.1 muestra los pasos generales de la metodología propuesta para alcanzar el objetivo de aumentar la resistencia al desgaste del material tratado superficialmente.

5.1.1 Pasos Generales de la Metodología Propuesta

1) Determinar las Variables Involucradas en el proceso (PTA) y sus niveles

Con base en la revisión de la literatura realizada en el apartado 3.1 del capítulo 3 se determinan las variables (entradas o parámetros del

proceso PTA) que tienen una influencia significativa sobre la resistencia al desgaste de la herramienta, así como sus niveles operacionales.

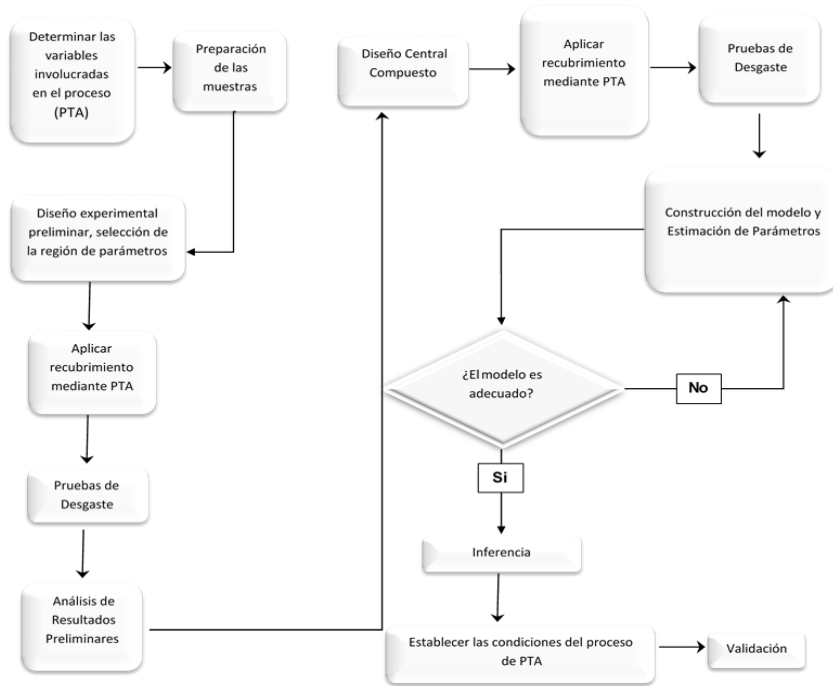


Figura 5.1. Metodología Propuesta. Fuente: Elaboración Propia

2) Preparación de las Muestras y diseño experimental preliminar

Se realiza un experimento piloto de diseño fraccionado 2^k , según los factores determinados en el paso número 1, para determinar el tamaño de la muestra y realizar una recolección de pruebas de forma ordenada, que represente los niveles establecidos para cada variable involucrada. Esta muestra será la probeta a utilizar para poder llevar a cabo las pruebas de desgaste. De forma tal que se puedan seguir los pasos siguientes de la metodología para obtener resultados que permitan

arribar a conclusiones parciales. Se procede también al corte del material para la preparación de las muestras, y una vez listas se prosigue con el tratamiento superficial (recubrimiento).

2.1) Preparación de las Muestras y Diseño Experimental

Con la información obtenida de las conclusiones parciales, se procede a la construcción de un diseño de experimentos central compuesto con combinaciones adicionales llamadas puntos axiales y n_c puntos centrales.

En la sección 4.5 del capítulo 4 se muestra la teoría relacionada con los diseños experimentales empleados.

3) Aplicar recubrimiento mediante PTA

Una vez determinados los niveles de los parámetros del proceso PTA y tomando como base el diseño experimental realizado en el paso anterior, se procederá a aplicar el recubrimiento a cada una de las muestras obtenidas, según lo estudiado en la sección 4.2 del Marco Teórico. El área donde variaran los niveles de los parámetros para la experimentación final, será determinada en función de los resultados obtenidos en la experimentación preliminar.

4) Pruebas de Desgaste

En este paso se someterá la muestra a la prueba de desgaste (Pin-on Disk) explicada en el apartado 4.3 del capítulo 4, para obtener el comportamiento de la capa del recubrimiento sometida a desgaste.

5) Construcción del modelo y estimación de parámetros

Una vez que se determina la tasa de desgaste, es necesario construir un modelo de regresión que sea capaz de representar el comportamiento adecuadamente, como caso particular la regresión cuantil a la cual se le dedica la sección 4.8 del presente trabajo. Para realizar la inferencia según la literatura se utilizará el método de estimación denominado Bootstrap.

6) Selección del modelo

Para la selección del modelo se desarrolla un análogo del estadístico R^2 utilizado en la regresión lineal (como se muestra en el apartado 4.8 del capítulo 4); comparando la suma de las distancias ponderadas para el modelo de interés con la suma en la que solo aparece el parámetro de intersección.

7) Validación del Modelo

En esta etapa se debe comprobar que los resultados aportados por el modelo para la variable de salida, no sean muy diferentes a los medidos en la realidad, y así conocer si el modelo resulta adecuado para obtener

conclusiones sobre el sistema real. Para ello resulta útil auxiliarse de la métrica a la que se hacía referencia en el apartado correspondiente a la regresión cuantil en el capítulo 4, la pseudo- R^2 .

8) Establecer las condiciones del proceso de PTA

Utilizar un método de optimización, que permita determinar las condiciones del proceso de PTA donde se mejore la resistencia al desgaste, utilizando como funciones objetivo los modelos construidos mediante regresión cuantil. Tal es el caso del algoritmo evolutivo multiobjetivo NSGA-II propuesto, en la sección 4.9, para aplicar en este caso puesto que no solo se optimizarán los parámetros de PTA, sino que es importante paralelamente analizar la dilución del recubrimiento, referida en el apartado 4.4, con el objetivo de obtener un valor adecuado también.

CAPÍTULO 6

6. Aplicación

En este capítulo se mostrará la aplicación de la metodología propuesta en un caso específico, mostrándose así todos los resultados obtenidos.

6.1 Determinación de las variables involucradas en el proceso (PTA) y sus niveles

Luego de una minuciosa revisión de literatura fueron seleccionados los parámetros para la experimentación inicial, y se representan en la tabla 6.3 (Molina, 2016) (Balasubramanian et al. , 2009)(Zamora, 2016).

Tabla 6.1. Parámetros del PTA

Parámetros	Valores
Tasa de alimentación de polvos	20 (g/min)
Flujo gas plasma	5 (l/min)
Altura del arco	1 (cm)
Flujo gas de arrastre	5 (l/min)
Flujo gas protección	12 (l/min)
Corriente de arco transferido	90 - 160 (A)
Velocidad de desplazamiento	10 - 20 (cm/min)

Se variarán solo aquellos que según la literatura ejercen mayor influencia, mientras que el resto se mantendrá fijo. Los factores escogidos y su dominio experimental se muestran en la tabla 6.2.

Para el diseño experimental se tendrá en cuenta también una variable discreta (recubrimiento), en cuyo caso es irrelevante qué nivel es el -1 y cuál es el $+1$. Esta notación codificada nos resultará útil para el factor discreto, ya que permitirá obtener una estimación numérica de su efecto a pesar de que el factor no toma valores numéricos.

Tabla 6.2. Factores y dominio experimental.

Factores	Dominio experimental	
	Nivel (-)	Nivel(+)
x_1 : Corriente de arco transferido (C) A	90	160
x_2 : Velocidad de desplazamiento (V) cm/min	10	20
x_3 : Recubrimiento (R)	Simple	Doble

6.2. Preparación de las muestras y diseño experimental preliminar

Se desarrolla una experimentación inicial, la cual conlleva inicialmente al análisis y preparación de los materiales con los que se va a trabajar.

6.2.1 Caracterización y preparación del material

Para la experimentación se utilizó como material base el Acero H13. Se cortaron muestras de 2 ½ pulgadas de diámetro y 2 cm de grosor, en una cortadora Dsicotom-10 Struers, con un disco de corte 40A25 de óxido de aluminio (250mm x 32 mm x 1.5 mm).

Se desbastaron y pulieron con el objetivo de lograr una superficie plana y libre de suciedades que puedan afectar la correcta deposición del recubrimiento.

El análisis químico del material, según sus proveedores, se muestra en la tabla 6.3.

Tabla 6.3. Composición química y dureza promedio del Acero H13

Material	C	Si	Mn	P	S	Cr	Mo	V	Dureza (BHN/HRC)
Acero H-13	0.37 0.43	0.90 1.20	0.30 0.50	≤ 0.030	≤ 0.030	4.80 5.50	1.20 1.50	0.90 1.10	241/23



Figura 6.1. Muestra de acero H13 (antes de pulir).

6.2.2 Material de aporte

Es una aleación de cobalto-cromo caracterizada por su alta resistencia a la corrosión, al desgaste y abrasión, además de buena resistencia mecánica, por lo que se puede utilizar para aumentar la resistencia al desgaste de los metales.

Tabla 6.4. Composición química y propiedades físicas del Stellite ®12 (Zamora , 2016).

Aleación	Co	Cr	W	C	Otros	Dureza (HV)	Densidad	Intervalo de Fusión (°C)
Stellite 12	Balance	27	7.5	1.4	Ni, Fe,	450	8.3 g/cm ³	1200
		32	9.5	1.7	Si, Mn	530		1365

Se desarrolla un experimento inicial 2^3 de fracción $\frac{1}{2}$, ya que se tienen 3 factores con dos niveles cada uno; es decir, solo se considera la mitad de las corridas experimentales. La Tabla 6.5 muestra el plan de experimentación que se obtiene al reemplazar los valores (-) y (+) por los valores de las variables reales.

El experimento se realiza con la intención de conseguir toda la información necesaria para el adecuado desarrollo de la investigación.

Según los resultados correspondientes a esta experimentación preliminar se sentarán las bases para realizar la experimentación final.

En esta primera fase se calcularán la dilución y el tiempo de desgaste del recubrimiento en cada una de las muestras. Ambas respuestas serán estudiadas, además de analizarse si existe alguna correlación entre ellas, pues de ahí se definirá el tipo de modelo a desarrollar.

Tabla 6.5. Matriz de experimentos preliminar

	Matriz de experimentos			Plan de experimentación		
	x_1	x_2	x_3	Corriente	Velocidad	Recubrimiento
1	-	-	-	90	10	1
2	+	-	-	160	10	1
3	-	+	-	90	20	1
4	+	+	-	160	20	1
5	-	-	+	90	10	2
6	+	-	+	160	10	2
7	-	+	+	90	20	2
8	+	+	+	160	20	2

6.3. Aplicación del recubrimiento mediante PTA

Para aplicar los recubrimientos se cuenta con un equipo de Plasma Transferido por Arco (PTA). Este equipo está integrado por una fuente de poder para arco piloto TransTig 1700, una fuente de poder para arco principal TransTig 4000, una tolva para alimentación de metal de aporte en polvo, un sistema de enfriamiento por refrigeración, una entrada para suministro de gas para arco piloto de protección y acarreo de polvo

(argón industrial), una entrada para suministro de gas para plasma (argón alta pureza 99.99%), una caja con controles de mando, una antorcha vertical y un sistema de desplazamiento horizontal con oscilación.

6.4. Pruebas de desgaste

Una vez realizado el recubrimiento con PTA, se desarrollan las pruebas de desgaste. Para ello se emplea un tribómetro Anton Paar modelo TBR, como se muestra en la figura 6.2; las condiciones para las pruebas se muestran en tabla 6.6.



Figura 6.2. Tribómetro Anton.

Tabla 6.6. Condiciones del ensayo de desgaste.

Parámetros	
Radio	2.50mm
Velocidad	10cm/s
Carga	8N

Se utiliza un perfilómetro marca Taylor Hobson, el cual tiene incorporado un “software” que permite realizar medidas de rugosidad superficial y de perfiles de las superficies. Este equipo está dotado de un palpador que consiste en una aguja acabada en una punta. Las medidas fueron hechas desplazando la punta sobre la superficie a medir. Las variaciones superficiales de la muestra causan desplazamiento vertical del palpador, con lo cual se genera una señal eléctrica correspondiente a los desplazamientos.

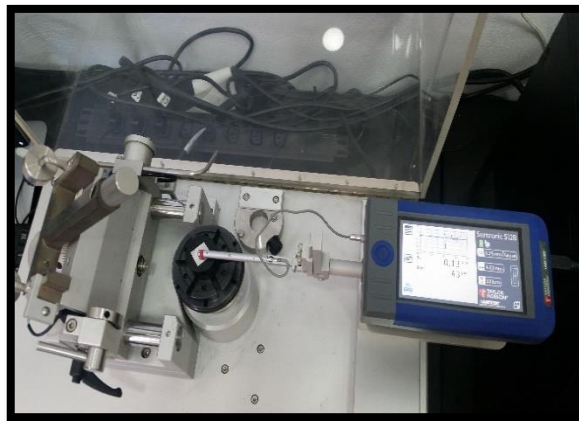


Figura 6.3. Perfilómetro conectado al tribómetro.

En este caso el perfilómetro ha sido utilizado para las medidas de profundidad y anchura de las huellas de desgaste, permitiendo obtener un promedio del área de las mismas.

Para determinar el volumen de la huella creada en la muestra se utiliza la ecuación (4.2) analizada en el capítulo 4. Una vez obtenido el volumen desgastado se procede a calcular la tasa de desgaste (K), a través de la ecuación (4.1).

6.5 Medición de la dilución

Para la medición de la dilución se pre-atacaron las muestras de la sección transversal con Nital al 2% durante 10 segundos y posteriormente se atacaron de 3 a 5 segundos con Berahas. Se utilizó el estereoscopio marca Nikon para tomar las imágenes de la zona de dilución. El porcentaje de dilución se midió en la sección transversal de las muestras utilizando la ecuación (4.3).

Tabla 6.7. Resultados de la medición del desgaste y la dilución.

	Matriz de experimentos			Plan de experimentación			Respuestas	
	x_1	x_2	x_3	Corriente	Velocidad de avance	Recubrimiento	y_1 (tasa de desgaste)	y_2 (dilución %)
1	-	-	-	90	10	1	6.020E-05	No existe
2	+	-	-	160	10	1	4.192E-04	52.91
3	-	+	-	90	20	1	6.380E-05	No existe
4	+	+	-	160	20	1	1.290E-04	55.62
5	-	-	+	90	10	2	1.700E-05	No existe
6	+	-	+	160	10	2	1.020E-04	32.74
7	-	+	+	90	20	2	1.79E-04	49.27
8	+	+	+	160	20	2	3.848E-04	22.61

La tabla 6.7 contienen los resultados obtenidos durante la experimentación piloto, después de calcular las tasas de desgaste y medir la dilución; evidenciándose que los recubrimientos que mostraron menor tasa de desgaste (muestras 1, 3 y principalmente la 5) presentaron un valor de dilución nulo, por lo tanto, serán descartadas.

Teniendo en cuenta esta información se representan los resultados en la figura 6.6 para definir la región de interés a partir de la cual será diseñado el experimento central compuesto.

Se deciden seleccionar las muestras 4 y 6 pues se encuentran entre las que presentaron menor tasa de desgaste, y además un nivel adecuado de dilución.

Por otra parte, se analiza la correlación existente entre el número de capas aplicadas en el recubrimiento y la dilución mediante la prueba de Pearson, obteniéndose un coeficiente de correlación de 0.05 y un valor P de 0.98; lo cual demuestra que no existe correlación entre la tasa de desgaste y la cantidad de capas.

Posteriormente al ser analizado este coeficiente para conocer la relación entre las capas aplicadas y la tasa de desgaste, el resultado nuevamente

demonstró que no existe correlación, obteniéndose un coeficiente de 0,02 y un valor P de 0.96. De esta manera se arriba a la conclusión de que tanto la dilución como la tasa de desgaste son dependientes del número de capas aplicadas en el recubrimiento.

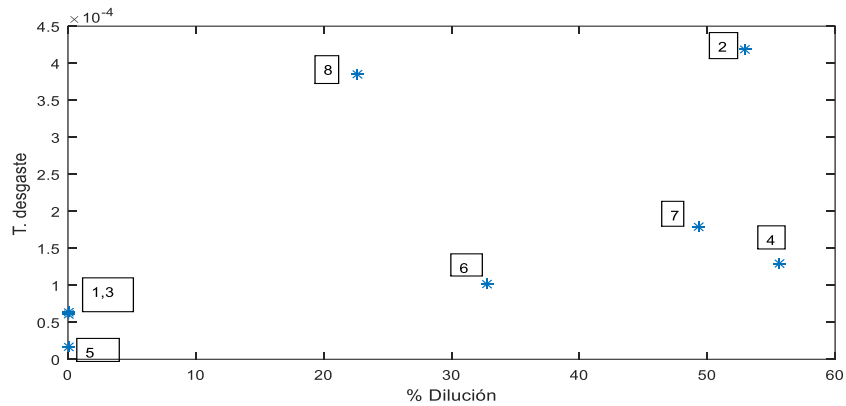


Figura 6.4. Análisis de la dilución y el desgaste

Basado en los resultados obtenidos, se procede a eliminar esta variable del experimento y plantear un diseño de experimento central compuesto, el cual es el indicado para realizar la elección de los niveles adecuados del proceso PTA. Las muestras en esta ocasión fueron sometidas a un tratamiento térmico previo, para garantizar una adecuada dureza. Los resultados obtenidos una vez completada la fase de experimentación se muestran en la tabla 6.8. Es necesario destacar que los ensayos de desgaste se realizaron bajo las mismas condiciones que en el experimento preliminar, al igual que para la medición del área de

desgaste, y posteriormente de la tasa de desgaste, se emplearon las mismas formulas; del mismo modo para la obtención de la dilución. En esta ocasión se obtuvieron todos los valores de la dilución, pudiéndose completar toda la recolección y análisis de los datos.

Se le realizo la prueba de correlación entre la tasa de desgaste y la dilución; donde se obtuvo un coeficiente de correlación de -0.029 y un valor P de 0.23; evidenciándose que no existe correlación entre la tasa de desgaste y el porcentaje de dilución. Lo cual indica que no será necesario utilizar un modelo de regresión multivariado; es decir, se ajustarán modelos independientes tanto para tasa de desgaste, como para dilución.

Tabla 6.8. Diseño de experimentos central compuesto.

Corrida	Matriz de experimentos		Plan de experimentación		Respuestas	
	x_1	x_2	Corriente	Velocidad de avance	y_1 (tasa de desgaste)	y_2 (dilución %)
1	-1	1	160.000	22.0711	2.65E-04	64.04
2	-1	1	145.858	15.0000	1.45E-04	51.97
3	0	1	160.000	15.0000	2.86E-04	52.00
4	1	1	170.000	20.0000	3.06E-04	40.18
5	1	1	150.000	20.0000	1.66E-04	57.80
6	-1	1	174.142	15.0000	1.35E-04	65.48
7	1	1	150.000	10.0000	2.40E-04	69.52
8	-1	1	160.000	7.9289	2.49E-04	69.32
9	0	1	160.000	15.0000	2.45E-04	55.23
10	0	1	160.000	15.0000	1.37E-04	66.15
11	1	1	170.000	10.0000	1.83E-04	63.36
12	0	1	160.000	15.0000	1.75E-04	53.27
13	0	1	160.000	15.0000	2.44E-04	62.79

6.6 Construcción del modelo

Considerando los datos presentados en la tabla 6.8; se obtuvieron los modelos para cada uno de los cuantiles, basados en las ecuaciones 4.21 y 4.22, para la tasa de desgaste y la dilución de manera independiente.

- **Estimación de parámetros**

Se encontraron los valores que minimizan la suma ponderada de las distancias entre valores ajustados (ecuación 4.25), obteniéndose los coeficientes correspondientes al ajuste de los modelos, los resultados se muestran en la tabla 6.9.

Tabla 6.9. Coeficientes de ajuste para los modelos.

<i>Cuantil</i>	<i>Coeficientes para la tasa de desgaste</i>			<i>Coeficientes para la dilución</i>		
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
0.01	7.687032e-05	-1.414227e-07	5.517155e-06	87.97874	-0.008454	-2.31804
0.99	3.061168e-04	-6.162832e-07	5.232566e-06	86.60013	-0.088535	-0.380393

- **Selección y validación del modelo**

La selección el modelo se realizó teniendo en cuenta la medida de ajuste, denominada pseudo- R^2 . Se seleccionaron los cuantiles .01 y .99 pues al ser los cuantiles extremos presentaron una mejor bondad de ajuste, obteniéndose una pseudo- R^2 de 0.969 y 0.971 para el caso del desgaste, respectivamente, y 0.974 para ambos cuantiles en el caso de la dilución. Estos valores indican un mejor ajuste en comparación con el resto de los cuantiles.

A continuación, se muestran las ecuaciones para cada uno de los modelos:

$$y(\text{desgaste})_i^{.01} = (7.687032e - 05) - (1.414227e - 07)x_1 + (5.517155e - 06)x_2 \quad (6.1)$$

$$y(\text{dilución})_i^{.01} = 87.97874 - 0.008454x_1 - 2.31804x_2 \quad (6.2)$$

$$y(\text{desgaste})_i^{.99} = (3.061168e - 04) - (6.162832e - 07)x_1 + (5.232566e - 06)x_2 \quad (6.3)$$

$$y(\text{dilución})_i^{.99} = 86.60013 - 0.088535x_1 - 0.380393x_2 \quad (6.4)$$

Observándose que al medir la bondad de ajuste utilizando la pseudo- R^2 , se pudo comprobar que los modelos representan el comportamiento de la resistencia al desgaste del acero H13 con recubrimiento de PTA de manera acertada, así como la dilución.

- **Inferencia**

Para la etapa de inferencia se utiliza el método de Bootstrap, el cual se introdujo previamente en la sección 4.9.4, del capítulo 4. Esta técnica de remuestreo permitió construir errores estándar, intervalos de confianza y pruebas de hipótesis con respecto a los modelos de regresión por cuantiles.

Para comprobar la significancia de los parámetros del PTA sobre el desgaste, se utilizó la prueba de Wald, descrita anteriormente en el capítulo 4, ecuación (4.27). Con esta prueba se puede determinar la significancia de cada uno de los factores que intervienen en el modelo; para ello se emplea el error estándar estimado.

Para cada uno de los coeficientes se obtiene se obtiene el valor de probabilidad (Valor P) para realizar la prueba de hipótesis. Considerando un α de 0.05 para un 95% de confianza; se puede comprobar en las tablas que los valores P menores de 0.05 pertenecen a los parámetros que afectan significativamente la respuesta, es decir, la tasa de desgaste.

Las tablas 6.10 y 6.11, muestran los valores de los intervalos estimados para los cuantiles .99 y .01 respectivamente.

Tabla 6.10. Inferencia para el cuantil .99

Coeficientes		Error estándar	L.I	L.S
$\hat{\beta}_0$	3.061168e-04	0.00090	-0.0015	0.0021
$\hat{\beta}_1$	-6.162832e-07	1.2E-08	-6.398E-07	-5.93E-07
$\hat{\beta}_2$	5.232566e-06	0.68200	-1.3367	0.5107

Tabla 6.11. Inferencia para el cuantil .01

Coeficientes		Error estándar	L.I	L.S
$\hat{\beta}_0$	7.687032e-05	0.00064	-0.0012	0.0013
$\hat{\beta}_1$	-1.414227e-07	2.00E-08	-1.81E-07	-1.02E-07
$\hat{\beta}_2$	5.517155e-06	0.00001	-1.41E-05	2.51E-05

Según los resultados mostrados, para un nivel de confianza del 95%, que se puede observar que existe un parámetro que es significativo, pues su intervalo no contiene al cero, y corresponde a la corriente de arco transferido. Desde la experimentación preliminar se pudo notar que efectivamente la variación entre los niveles de este factor influía sobre la tasa de desgaste y la dilución principalmente.

6.7 Establecer las condiciones del proceso de PTA

Para encontrar las condiciones óptimas del proceso de PTA, fue necesario tener en cuenta no solo que se alcanzara una menor tasa de desgaste, sino también que mantuviera un adecuado nivel de dilución, para el cual se fijó un valor del 50%. Dadas las características de los modelos ajustados, se utilizó el optimizador multiobjetivo NSGA II, del cual se hace mención en la sección 4.10 del presente trabajo. Se tomaron como funciones objetivo las ecuaciones (6.1), (6.2), (6.3) y (6.4), y las restricciones de las variables involucradas en el proceso de PTA, las cuales se muestran en la tabla 6.12:

Tabla 6.12. Restricciones del problema de optimización.

Variables del PTA	Inferior	Superior
Corriente	145.858	174.142
Velocidad de avance	7.9289	22.0711

Los resultados obtenidos luego de procesar los datos mediante el algoritmo NSGA II, se muestran en la tabla 6.13. Se encontraron dos combinaciones de niveles que demostraron tener los mejores resultados en base al objetivo del problema, obtener una mayor resistencia al desgaste manteniendo una adecuada dilución, 50%.

Tabla 6.13. Resultados de la optimización.

	Niveles óptimos		Intervalo de resultados esperados	
	Corriente	Velocidad de avance	Tasa de desgaste	Dilución (%)
1	157.856	15.808	1.4190380e-04 — 2.9143771e-04	50 — 66.611
2	171.792	14.995	1.3545246e-04 — 2.7860266e-04	51.766 — 65.686

Para validar las condiciones óptimas del proceso de PTA obtenidas se realiza nuevamente el proceso de recubrimiento en dos muestras, la tasa de desgaste y la dilución resultante para cada una de las combinaciones de niveles encontradas para los parámetros del proceso se muestran en la tabla 6.14.

Tabla 6.14. Resultados de la validación en campo

	Niveles óptimos		Intervalo de resultados esperados	
	Corriente	Velocidad de avance	Tasa de desgaste	Dilución (%)
1	157.856	15.808	2.62365e-04	65.9
2	171.792	14.995	2.49167e-04	64.6

Se pudo comprobar que efectivamente se obtuvieron los valores dentro del intervalo de resultados esperados.

CAPÍTULO 7

7. Conclusiones

Durante el desarrollo del trabajo, y con la aplicación de la metodología propuesta, se logró representar los datos del desgaste a través de la regresión cuantil. Se construyeron modelos que condujeron a estimaciones acertadas y además pudieron ser optimizados, obteniéndose así valores que indican las condiciones óptimas en las que se maximiza la resistencia al desgaste del recubrimiento mediante PTA sobre el acero H13 (**comprobación de la hipótesis general**). De esta manera, también se le dio cumplimiento al **objetivo general** de la investigación.

Respecto a los objetivos específicos, al estar estrechamente vinculados con la metodología propuesta, se les dio cumplimiento en la medida que era aplicada la misma. Las variables involucradas en el proceso de PTA fueron determinadas a través de una profunda revisión de la literatura, donde se identificaron aquellas que tienen un efecto más significativo en cuanto a la resistencia al desgaste, las cuales fueron la corriente de arco transferido y la velocidad de avance. A través de los diseños de experimentos empleados se organizaron y ejecutaron las pruebas de desgaste, así como la recolección de datos. Se logró seleccionar los

modelos que representaran adecuadamente los datos teniendo en cuenta la variabilidad, dada la robustez que caracteriza a la regresión cuantil. Mediante el análisis de la pseudo- R^2 como medida de bondad de ajuste se seleccionaron aquellos modelos pertenecientes a los cuantiles que verdaderamente eran representativos (0.01 y 0.99), y que presentaron valores de pseudo- R^2 muy cercanos a 1. Finalmente, en cumplimiento del último de los objetivos específicos, se empleó un algoritmo de optimización multiobjetivo, el NSGA II (**respuesta a la pregunta de investigación 5**), que permitió determinar los niveles donde se maximice la resistencia al desgaste del recubrimiento y se obtenga una adecuada dilución, considerando como función objetivo el modelo de regresión cuantil para ambos casos (**respuesta a la pregunta de investigación 3**).

Para obtener los valores del desgaste se utilizó como indicador la tasa de desgaste, por considerarse en la literatura como un parámetro cuantitativo universal para medir el desgaste, y uno de los más utilizados por brindar valores acertados en dependencia del volumen perdido (**respuesta a la pregunta de investigación 1**).

En respuesta a la **segunda pregunta de investigación**, primeramente, se ejecutó un experimento piloto mediante un diseño 2^k de fracción 1/2, donde la región experimental para cada una de las variables que

intervinieron en el recubrimiento fue seleccionada a partir de la revisión de literatura. Basados en los resultados obtenidos en esta primera etapa de la investigación se diseña un experimento central compuesto, donde se establece la región experimental para cada uno de los factores, corriente de arco transferido (150- 170 A) y velocidad de avance (10-20 cm/min).

A través del estudio de la regresión cuantil, se encontró la manera de estimar los parámetros del modelo, mediante la minimización de la suma ponderada de los residuales (**respuesta a la pregunta de investigación 4**).

Finalmente se desarrolló la validación en campo con los parámetros propuestos en la optimización, con lo que se pudo autentificar los resultados obtenidos en la presente investigación.

BIBLIOGRAFÍA

- Aguirre Sánchez, Á. (2012). Optimización de los parámetros del proceso de plasma trasferido por arco (PTA) utilizado en la recuperación de componentes de acero grado herramienta D2. Tesis de Maestría. Saltillo, Coahuila
- Allen D.M. (1971). Mean square error of prediction as a criterion for selecting variables. *Technometrics* 13, 469-475.
- Alvarado, C., Herazo, I., Ardila, C., & Donoso, Y. (2005). Aplicación de NSGA-II y SPEA-II para la optimización multiobjetivo de redes multicast. *Ingeniería& Desarrollo* (17), 29-53.
From <https://www.researchgate.net/publication/28200496>.
- Ayala, M., Borges, R., & Colmenares L, G. (2007). Verificación de los supuestos del modelo Cox. Caso de estudio: Banca comercial venezolana. *Economía*(23), 27-43. From <http://www.redalyc.org/articulo.oa?id=195617571003>
- Balasubramanian, V., Lakshminarayanan , A., Varahamoorthy , R., & Babu, S. (2008). Understanding the Parameters Controlling Plasma Transferred Arc Hardfacing Using Response Surface Methodology. *Materials and Manufacturing Processes*, 674–682. doi:10.1080/15560350802316744
- Balasubrarnanian, V., Lakshminarayanan , A., Varahamoorthy , R., & Babu, S. (2009). Application of Response Surface Methodolody to Prediction of Dilution in Plasma Transferred Arc Hardfacing of Stainless Steel on Carbon Steel. *Journal of iron and steel research, International*, 44-53.
- Box, G., Hunter, W., & Hunter, J. (2005). *Statistics for experimenters: design, innovation and discovery*. John Wiley & sons, Hoboken, New Jersey.
- Brett, T., & Wolfe, B. (2010). Homogeneity of Metal Matriz Composites Deposuted by Plasma Transferred Arc Welding. Edmonton: Spring. From <https://era.library.ualberta.ca/items/aca5df77-4de8-476e-8a40-0017a110e154>

- Buchinsky, M. (1998). "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research. *Journal Of Human*, 88-126.
- Byungtae, S., Jungsik, N., Taewook, L., & Young, J. (2017). Adaptative robust regression with continuos Gaussian scale mixture errors. *Journal of the Korean Statistical Society*, 113-125. From <https://www.sciencedirect.com/science/article/pii/S1226319216300345?via%3Dihub>.
- Cabot, P., Moglioni, A., & Carella, E. (2003). Soldadura por Fricción Agitación. *Materia*, 187-195.
- Carmonab, G., & Correa, J. C. (2015, Diciembre). Comparison of Gini Regression with OLS Regression and other Robust Linear Regression. 8(2), 129-161.
- Carrasco, J., Berdugo, I., Ospina, R., & Limy, U. (2013). Optimización del diseño de herramienta con pin cónico roscado para soldadura por fricción-agitación. *Dialnet* (135-144).
- Cervantes Hernández, P., Flores Gómez, A., & Sánchez Meraz, B. (2005). Mínimos cuadrados versus verosimilitud. *Ciencia y Mar*, 41-45.
- Chavarría Mayorga, J. (2014). Predicción de volatilidad mediante regresión por cuantiles. *Revista Científica de la UNAN - León*, 72-93.
- Cochran, W., & Cox, G. (1990). *Diseños experimentales*. Mexico: Trillas, From <https://www.urbe.edu/UDWLibrary/InfoBook.do?id=5068>
- Conde Amboage, M. (2013). Contraste de Bondad de Ajuste de Modelo de Regresión Cuantil. *Universidad Santiago de Compostela*.
- Correa Flórez, C., Bolaños, R., & Molina Cabreras, A. (2008). Algoritmo multiobjetivo NSGA-II aplicado al problema de la mochila. *Scientia et Technica*(39), 206-211.

- Cruz Crespo, A., Fernandez Fuentes, R., & Scotti, A. (2011). Efecto sobre la dilución de la Granulometría de la ferroaleación en el alma de electrodos tubulares revestidos bajo la influencia de la composición del revestimiento. *Soldag. insp. São Paulo*, 16(1), 79-85.
- Cruz, E. (2006). Aplicación de un Modelo Factorial de Experimentación en un Ingenio Azucarero del Valle del Cauca. *Revista Scientia et Technica*, 313-318.
- Davino, C., Furno, M., & Vistocco, D. (2014). *Quantile Regression Theory and Applications*. John Wiley & Sons, Ltd
- Deb, K., Agrawal, S., Pratap, A., & Meyarivan, T. (2000). A fast and elitist multiobjective genetic algorithm: NSGA-II. (I. I. Technology, Ed.) Kanpur Genetic Algorithms Laboratory, 15.
- Díaz, A., & De Vasconcelos, J. (2002, Marzo). Multiobjective Genetic Algorithms. *IEEE Transactions on Magnetics*, 38(3), 1133-1136.
- Duardo Toste, A. (2015). Optimización multiobjetivo de mallas de tierra de subestaciones. Santa CLara: Facultad de Ingeniería Eléctrica.
- Echavarría Figueroa, C. J. (2016). Optimización de parámetros y análisis de resistencia al desgaste en recubrimientos base Ni sobre acero A 36 mediante Plasma Transferido por Arco (PTA). Comimsa, Saltillo, Coahuila.
- Echevarría, C. (2016, Noviembre). Optimización de parámetros y análisis de resistencia al desgaste en recubrimientos base Ni sobre acero A36 mediante plasma transferido por arco (PTA). Coahuila, Saltillo.
- Elkan, C. (2011). *Evaluating Classifiers*. University of California, San Diego.
- Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2013). *Regression Models, Methods and Applications*. New York: Springer.

- Ferré, J., & Rius, X. (2004). Diseño factorial completo 2k. Técnicas de Laboratorio 292, 430-434. From <http://www.quimica.urv.es/quimio>
- Fox, J., & Weisberg, S. (2013). Car: companion to applied regression. From <http://cran.r-project.org/web/packages/car/index.html>.
- Geisser S. (1993). Predictive Inference: An Introduction, Monographs on statistics and applied probability. Chapman & Hall, First Edition.
- González Gonzalez, D. (2010). Incorporación de los Efectos que el Medio Ambiente Operacional tiene sobre la Confiabilidad de un sistema. Tesis de Doctorado. Saltillo, Coahuila
- Gutiérrez Pulido, H., & de la Vara Salazar, R. (2004). Control estadístico de calidad y seis sigma. Mexico: McGraw-Hill Interamericana.
- Hao, L., & Naiman, D. (1949). Quantile Regression. SAGE Publications.
- Hernández, C. A., Castilla, G., & Mancilla, J. E. (2016). Un algoritmo multi-objetivo NSGA-II para la programación de los pasos de laminación en caliente de aceros. Research in Computing Science, 65-80.
- Holland, J. (1988). The dynamics of searchers directed by genetic algorithms. (Y. Lee, Ed.) Evolution, learning, and cognition, Teaneck, N.J. World Scientific.
- Jiménez Careaga, M. (2015). Superficies de Respuesta mediante un Diseño Central Compuesto. Varianza, 31-36.
- Jiménez, J., Alamilla, N., & López, M. (2009). Pruebas de vida acelerada. Temas de Ciencia y Tecnología (Vol. 13).
- Koenker, R. (2005). Quantile Regression. Cambridge, UK: Cambridge University Press.
- Lee, B.-J., & Lee, M. (2006). Quantile Regression Analysis of Wage Determinants in the Korean Labor Market. 1-31.

- Leng, C., & Tong, X. (2013). A quantile regression estimator for censored data. *ISI/BS in Bernoulli*, 344-361.
- Luengo López, O. (2008). Estudio tribológico de fricción y desgaste de capas de aleación NiP. Tesis de licenciatura, Escuela Técnica Superior de Ingeniería Industrial de Barcelona, Barcelona.
- Marchena, W., & Carlos, O. (2007). Optimización y la Programación Lineal: Una Introducción.
- Martínez Estudillo, A. (2005). Modelos de Regresión Basados en Redes Neuronales de Unidades Productos Diseñadas y Entrenadas Mediante Algoritmos de Optimización Híbrida. Aplicaciones. Tesis Doctoral. Universidad de Granada, España.
- Medina Varela, P. D., & Lopez Reyes, A. M. (2011). Análisis crítico del diseño factorial 2k sobre casos aplicados. *Scientia Et Technica*, pp. 101-106.
- Mohaideen Ferozhkhana, M., Duraiselvam, M., Ganesh kumar, K., & Ravibharath, R. (2015). Plasma Transferred Arc Welding of Stellite 6 Alloy on Stainless Steel for Wear Resistance . *Procedia Technology*, 1305 – 1311.
- Molina Claros, J. (2016). Efecto de la adición de nanopartículas de carburo de tungsteno en la microestructura y resistencia al desgaste de un recubrimiento basa níquel obtenido mediante el rproceso de plasma transferido por arco (PTA). Saltillo.
- Montgomery, D. (2008). Introduction to regression analysis. Mexico: Continental.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2006). Introducción al análisis de Regresión Lineal (Tercera ed.). Mexico: Continental.
- Peñuela Meneses, C. A., & Granada Echeverri, M. (2007). Optimización multiobjetivos usando un algoritmo genético y un operador elitista basado en un ordenamiento no-dominado (NSGA-II). *Scientia et Technica*, 175-180.

- Pérez, S. R. (2010). [www.monografias.com](http://www.monografias.com/trabajos93/motores-electricos/motores-electricos.shtml#ixzz2WggRMeNG). Retrieved junio 15, 2013 from <http://www.monografias.com/trabajos93/motores-electricos/motores-electricos.shtml#ixzz2WggRMeNG>
- Phatak, A., & Robinson, G. (2005). Understanding and modelling variability: Practitioners Perspectives. From <http://www.stat.auckland.ac.nz/~iase/publications.php?show=13>
- Portnoy, S. (2003). Censored Regression Quantiles. *Journal of the American Statistical Association*, 1001-1012.
- Rabinowicz, E. (1995). *Friction and Wear of Materials*. John Wiley and Sons, New York.
- Sánchez, B., & Vicéns, J. (2012, Abril). Regresión Cuantílica: Estimación y Contrastes. From <http://www.uam.es/klein/gauss>
- Sun, X., Peng, L., & Amita, M. (2016). Quantile Regression Analysis of Censored Longitudinal Data with Irregular Outcome-Dependent Follow -Up. *Biometrics*, 64-73.
- Surendra P, S., Ramoni P, J., Torres R, E., & Orlandoni M, G. (2010). Modelado del Índice de Capacidad Profesional de profesores universitarios por regresión cuantil: El caso de la Universidad de los Andes. *Economía*, XXXV, 209-225.
- Treviño, Á. (2004, Junio). Desarrollo de una metodología para pruebas de erosión. Nuevo León, México.
- Urbano, L. O., Gordillo, M., & Franco, F. (2013). Optimización de Soldadura por Fricción - Agitación de la aleación de Aluminio AA 6261 - T5 mediante la aplicación de Metodología de Superficie de Respuesta. *LatinAm.Metal.Mat.*
- Vicéns Otero, J., & Sánchez Reyes, B. (2012). Regresión Cuantílica: Estimación (Vol. No.21).

- Villada Fernando, C. R. (2007). Diagnostico de fallas en motores de inducción mediante la aplicación de redes neuronales artificiales. Información Tecnológica.
- Wu, Y., & Yin, G. (2013). Cure Rate Quantile Regression of Censored Data With a Survival Fraction. *Journal of the American Statistical Association Publication Details*, 1517-1531.
- Wu, Y., & Yin, G. (2017). Multiple Imputation for Cure Rate Quantile Regression with Censored Data. *Biometrics*, 94-103.
- Zamora López, A. V. (2016). Efecto de la adición de nanopartículas de WC y TIC en la modificación superficial del acero grado herramienta H 13 mediante el proceso de Plasma Transferido por Arco (PTA.). COMIMSA, Saltillo, Coahuila.
- Zeng, W., Wu, H., & Zhang, J. (2006). Effect of tool wear on microstructure, mechanical properties and acoustic emission of friction stir welded 6061 Al alloys. *Acta Metall Sinica*, 9-19.
- Zitzler, E., Deb, K., & Thiele, L. (2000). Comparison of multiobjective evolutionary algorithms: empirical results. *Evol Comput* 8, 173-195.