



**CORPORACIÓN MEXICANA DE INVESTIGACIÓN EN MATERIALES**

**DIVISIÓN DE ESTUDIOS DE POSGRADO**



**"Análisis multivariado del proceso de fundición por molde permanente  
mediante una red neuronal de Función de Base Radial"**

**TESIS**

**Para obtener el Grado Académico de**

**Maestro en Ciencia y Tecnología en Ingeniería Industrial y de Manufactura**

**Presenta:**

**Homero de Jesús de León Delgado**

**Saltillo, Coahuila 30 de noviembre de 2016**

***"Análisis multivariado del proceso de fundición por molde permanente  
mediante una red neuronal de Función de Base Radial"***

*Por  
Homero de Jesús de León Delgado*

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*

*Maestría en Ciencia y Tecnología  
**En Ingeniería Industrial y de Manufactura***

**Programa Interinstitucional en Ciencia y Tecnología COMIMSA / CONACyT**

**Saltillo, Coahuila noviembre de 2016**

## RESUMEN

En la actualidad las empresas de manufactura han visto incrementada la dificultad en la toma de decisiones de sus procesos. Debido a esta dificultad se utilizan herramientas como el modelado de procesos, para tomar decisiones en base al comportamiento del modelo.

La red neuronal de Función de Base Radial es un sistema inteligente aplicado al modelado en la industria de manufactura. Sin embargo, la mayoría de los procesos presentan varias salidas dificultando el análisis, debido a esto, es común encontrar trabajos donde se ignora la relación entre variables de salida, provocando diversos problemas al tomar decisiones en base a este tipo de análisis. El presente trabajo propone el uso de las redes neuronales de Función de Base Radial para la predicción en procesos de manufactura con varias variables de respuesta y su análisis por medio de herramientas estadísticas multivariadas como el MANOVA, tomando en cuenta las condiciones que se deben cumplir para aplicar este tipo de análisis. Esto para conocer que tan bien explica el modelo la variabilidad del proceso y qué variable es más importante.

La aplicación de la metodología propuesta fue en un proceso de fundición en molde permanente y por gravedad, donde se obtuvo una buena predicción en base a métricas estadísticas multivariadas y se encontró que de las variables basculamiento, temperatura del metal y temperatura del molde, esta última fue la de mayor influencia en el proceso.

**Palabras clave:** *Red Neuronal, Función de Base Radial, MANOVA*

## Tabla de contenido

Capítulo 1: Introducción .....	1
Capítulo 2: Planteamiento del problema .....	7
2.1 Descripción del problema .....	7
2.2 Preguntas de investigación.....	15
2.3 Hipótesis.....	16
2.4 Objetivos.....	16
2.4.1 Objetivo general .....	16
2.4.2 Objetivos específicos.....	16
2.5 Justificación .....	17
2.6 Delimitaciones .....	19
Capítulo 3: Estado del arte .....	20
3.1 Conclusiones del estado del arte.....	27
Capítulo 4: Marco teórico .....	29
4.1 La red neuronal artificial .....	29
4.2 Funciones de activación .....	31
4.3 Funciones de Base Radial.....	31
4.4 Interpolación por una red de Función de Base Radial .....	32
4.5 Teorema de Micchelli.....	35
4.6 Funciones Gaussianas Multivariadas .....	36
4.7 Redes FBR generalizadas.....	37
4.8 Estimación de los pesos de la red por mínimos cuadrados.....	38
4.9 Estimación de los centros de la red por Algoritmo Genético .....	40
4.10 Distancia de Mahalanobis.....	42
4.11 Análisis Multivariado de Varianza (MANOVA) .....	42
4.12 Elipsoides de confianza e intervalos.....	52
4.13 Medidas de asociación y validación de la predicción.....	53
Capítulo 5: Metodología.....	56
Capítulo 6: Aplicación .....	62
6.1 Descripción del proceso de manufactura.....	62

6.2 Análisis de datos.....	63
6.3 Prueba de correlación.....	69
6.4 Modelación .....	72
6.5 Cumplimiento de supuestos .....	74
6.6 Análisis de varianza multivariado .....	77
6.7 Validación .....	79
Capítulo 7: Conclusiones .....	85
Bibliografía.....	89

## Lista de Tablas

Tabla 2.1 Datos proceso EDM.....	8
Tabla 2.2 Supuestos con BP.....	9
Tabla 2.3 Supuestos con FBR .....	10
Tabla 2.4 Correlación proceso EDM.....	11
Tabla 2.5 Valores de probabilidad proceso EDM.....	11
Tabla 4.1 Construcción de la tabla MANOVA .....	51
Tabla 6.1 Datos del proceso .....	64
Tabla 6.2 Puntos centrales de $y_1$ .....	65
Tabla 6.3 Medias de cada cavidad de $y_1$ .....	66
Tabla 6.4 Puntos centrales de $y_2$ .....	67
Tabla 6.5 Medias de cada cavidad de $y_2$ .....	67
Tabla 6.6 Puntos centrales de $y_3$ .....	68
Tabla 6.7 Medias de cada cavidad de $y_3$ .....	68
Tabla 6.8 Variables de respuesta .....	70
Tabla 6.9 Coeficientes de correlación de Pearson.....	70
Tabla 6.10 Coeficientes de correlación de Kendall .....	71
Tabla 6.11 Centros generados por algoritmo genético .....	72
Tabla 6.12 Pesos generados por mínimos cuadrados.....	73
Tabla 6.13 MANOVA para significancia del modelo .....	78
Tabla 6.14 Medidas de asociación y predicción multivariadas .....	78
Tabla 6.15 MANOVA para subconjuntos de $x$ .....	79
Tabla 6.16 Predicciones con la red FBR.....	80
Tabla 6.17 Intervalos de confianza para los pesos .....	81
Tabla 6.18 Intervalos de confianza para las predicciones .....	82

## Lista de Figuras

Figura 1.1 Esquema de un proceso.....	1
Figura 4.1 Estructura de una red neuronal .....	30
Figura 4.2 Estructura de la red FBR .....	38
Figura 5.1 Metodología general .....	56
Figura 5.2 Análisis de datos.....	57
Figura 5.3 Correlación no significativa .....	60
Figura 5.4 Correlación significativa .....	60
Figura 5.5 Validación .....	61
Figura 6.1 Metodología propuesta .....	62
Figura 6.2 Pieza fabricada .....	64
Figura 6.3 Gráfica de caja para $y_1$ .....	66
Figura 6.4 Gráfica de caja para $y_2$ .....	67
Figura 6.5 Gráfica de caja para $y_3$ .....	69
Figura 6.6 Intervalos de confianza para Peso.....	83
Figura 6.7 Intervalos de confianza para SPC 2 .....	83
Figura 6.8 Intervalos de confianza para SPC 3 .....	84

# Capítulo 1

## Introducción

En la actualidad las empresas de manufactura han visto incrementada la dificultad en sus procesos de toma de decisiones, debido a los rápidos cambios en los métodos de diseño y la demanda de productos de calidad (Ahilan et al., 2013). Esta tendencia de constante cambio y la necesidad de mantenerse a la vanguardia tecnológica ha propiciado la necesidad de tener procesos cada vez más eficientes.

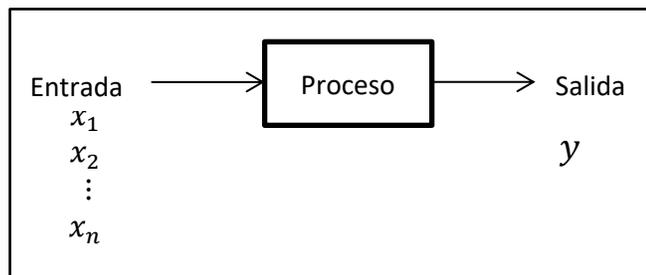


Figura 1.1 Esquema de un proceso

Los procesos de manufactura son un conjunto de actividades donde se realizan modificaciones en las características de un producto. Los más importantes son: maquinado, fundición, soldadura y conformado (Chandrasekaran et al. 2010), los cuales describen un comportamiento sistémico (Figura 1.1) con sus elementos de entrada (energía requerida, ciclos, potencia, etc.), una fase de transformación de las entradas y las salidas del proceso (características de calidad, energía utilizada, etc.)

Dado que aumentar la eficiencia en los procesos es uno de los objetivos principales de la industria manufacturera, resulta útil apoyarse en el uso de diversas metodologías y herramientas para el control de la producción como lo es el modelado, ya que esto permite representar el proceso y tomar decisiones en base al comportamiento explicado por el modelo.

Gracias al modelado el proceso de optimización se facilita, ya que normalmente el ajuste de parámetros y variables se da de manera aleatoria o por la práctica de “prueba y error”, esto genera gastos en todos los sentidos; costos por piezas defectuosas, tiempo de producción o gastos en materiales (Benyounis y Olabi, 2008). Por lo tanto, se tiene la necesidad de mejorar la predicción y optimización de los procesos de manufactura, y sobre todo en el manejo de parámetros y variables.

Uno de los procesos de manufactura más importante es la fundición (Chandrasekaran et al., 2010), el cual consiste en fundir un material para vaciarlo en estado líquido a un molde para obtener la forma deseada. Existen diferentes tipos de fundición: con molde en arena, molde permanente, cera perdida, inyección, etc. En el caso de estudio se trabaja con un proceso de fundición en molde permanente y por gravedad, donde es muy importante seleccionar los mejores parámetros y analizar los efectos entre variables controlables para el aseguramiento de la calidad. Es común encontrar estos parámetros por prueba y error utilizando herramientas de modelado para encontrar las condiciones adecuadas del proceso.

Existen diferentes técnicas tanto para modelar, como para optimizar un proceso. Las herramientas clásicas son estadísticas como la regresión (Cevallos, 2008), pero también existen métodos basados en sistemas inteligentes, los cuales simulan o tratan de imitar el comportamiento de los

sistemas biológicos humanos, tales como redes neuronales, lógica difusa, algoritmos genéticos y combinaciones en sistemas híbridos. Fueron desarrollados para la resolución de problemas y la toma de decisiones en problemas complejos (Yen & Langari, 2000). Dentro de las redes neuronales existen diferentes tipos; máquinas de vectores de soporte, teoría de resonancia adaptativa, retro-propagación, redes de función de base radial, etc. Estas dos últimas tienen como aplicaciones principales, en diversas áreas de la ingeniería, la clasificación y la predicción. Son nombradas “aproximadoras universales” (Haykin, 1999, Cevallos, 2008) por su capacidad de aproximar los valores de entrada a las salidas deseadas gracias a las funciones de base radial, además pueden aplicarse en problemas no lineales y dado que los procesos de manufactura tienen comúnmente un carácter no lineal (Cevallos, 2008), resulta una herramienta útil para analizar los procesos de manufactura y optimizarlos.

Por su parte, el modelado permite explicar las salidas  $y$  del proceso en base a las variables de entrada  $x$ . Cada especificación de entrada utilizada en el proceso afecta de manera directa a su resultado, tales datos se representan como variables de entrada en una matriz  $X$ .

$$[X] = [x_1, x_2, \dots, x_p]$$

Donde  $x_1, x_2, \dots, x_p$  son cada una de las  $p$  variables representativas del proceso.

Mientras que las salidas de un proceso univariado se representan en un vector  $[y]$ , y en un proceso multivariado en una matriz de respuestas  $[Y]$ .

$$[Y] = [y_1, y_2, \dots, y_q]$$

Donde  $y_1, y_2, \dots, y_q$  son las  $q$  variables de respuesta del proceso y pueden ser características de calidad resultante o condiciones finales del proceso.

Es común que los procesos de manufactura generen más de una respuesta. Por ejemplo, un proceso de soldadura genera una profundidad en la penetración de la soldadura, pero también genera un diámetro de la misma, un maquinado por descargas eléctricas (EDM) genera un rango de material removido y cierta rugosidad en la superficie (Mollah et al. 2008), un maquinado con CNC también genera rugosidad en la pieza y resulta útil medir la energía consumida por el proceso (Ahilan et al. 2013). El proceso de fundición no es la excepción, como en los casos de Schafföner et al., (2016), Timelli et al., (2016) y Patel et al., (2016), donde utilizan herramientas estadísticas para analizar procesos de fundición con varias respuestas.

Para analizar procesos de manufactura con más de una respuesta como los vistos anteriormente, es común que se utilicen modelos univariados para cada respuesta del proceso. También, se han desarrollado metodologías de redes neuronales para analizar todas las respuestas considerándolas independientes entre sí, como las desarrolladas por los autores Du et al., (2010), o Pérez et al., (2014), entre otros. Sin embargo, considerar las respuestas independientes genera problemáticas a la hora de optimizar el proceso. Por ejemplo, el modelo puede ajustarse adecuadamente solo para una respuesta, y encontrar parámetros óptimos para una respuesta y diferentes en otra. Por las características del proceso posiblemente los parámetros deben ser los mismos para todas las características de calidad y respuestas resultantes. Este caso se daría al tener respuestas correlacionadas en el proceso.

Al analizar el proceso surgen preguntas como:

¿Qué variable del proceso afecta más a las respuestas de manera conjunta?

¿Cómo sabemos que el modelo explica el proceso de forma adecuada?

Para responder esto se sugiere el uso del análisis de varianza, sin embargo, existen ciertos supuestos estadísticos que se deben tomar en cuenta para llegar a conclusiones acertadas por medio del análisis de varianza, pero:

¿Cómo se verían afectados estos supuestos al existir correlación en múltiples salidas?

Al considerar varias respuestas es posible preguntarse:

¿Cómo demostrar la correlación entre las respuestas?

Como se mencionó anteriormente, las redes neuronales artificiales cuentan con diferentes tipos, de las cuales las de Retro-Propagación (RP) y las Funciones de Base Radial (FBR) tienen especial importancia en la industria ya que se les conoce como aproximadores universales (Haykin, 1999), y pueden ser usadas en aplicaciones de predicción y clasificación. Su uso en predicción dentro de los procesos de manufactura es recomendable gracias a su capacidad de trabajar en procesos complejos. La determinación de los pesos sinápticos en la red resulta clave para lograr las salidas deseadas, ya que una vez determinando los pesos, las entradas pueden cambiar para encontrar los parámetros óptimos. Para esto, resulta útil la combinación de la red con herramientas estadísticas como la estimación por mínimos cuadrados, logrando con esto que el error a la salida de la red sea mínimo mejorando la calidad de la predicción. Aquí surgen otras preguntas como:

¿En qué afectan las múltiples salidas en la estimación de los pesos de la red?

El presente trabajo está estructurado en siete capítulos con el fin de responder las cuestiones planteadas a lo largo de este capítulo uno. En el capítulo dos se plantea la problemática de utilizar el análisis de varianza sin verificar los supuestos necesarios para este análisis, además de las consecuencias de ignorar la relación entre variables de salida, se plantean los objetivos, las preguntas de investigación, las hipótesis y justificación del proyecto.

En el capítulo tres se desarrolla el estado del arte referente a las aplicaciones de diferentes tipos de trabajos en diferentes áreas y procesos multirespuesta donde se utilicen redes neuronales principalmente en predicción.

El capítulo cuatro muestra una introducción al marco teórico detallando las herramientas y conceptos utilizados para el desarrollo del modelo. El capítulo cinco detalla la metodología planteada para analizar procesos de manufactura con una o varias respuestas, mediante la red de Función de Base Radial y el Análisis de Varianza.

El capítulo seis muestra la aplicación y resultados obtenidos en el proceso de fundición en molde permanente y por gravedad, mediante la metodología propuesta. Finalmente, en el capítulo 7 se presentan las conclusiones de la investigación y comentarios sobre el trabajo futuro.

## **Capítulo 2**

### **Planteamiento del problema**

En el presente capítulo se presentará el problema principal y su entorno, objetivos e hipótesis.

#### **2.1 Descripción del problema**

Dadas las características mencionadas en el capítulo uno, es importante verificar si el análisis del modelo puede realizarse bajo las mismas condiciones si se tienen varias respuestas en el proceso. En general, un modelo con varias respuestas puede dividirse en dos categorías: sistemas con respuestas independientes y sistemas con respuestas correlacionadas (Bashiri et al. 2014). En la primera categoría puede considerarse cada respuesta independiente y es posible verificar los efectos de algún cambio de nivel de los factores de control en cada respuesta sin considerar otras. Bajo este método es lógico pensar que el análisis y la optimización de estos sistemas son más simples. Por otro lado, en procesos con respuestas correlacionadas, no es posible analizar cada respuesta independientemente ya que conduciría a interpretaciones erróneas. Debido a la dificultad del análisis y optimización de procesos con respuestas correlacionadas, es común encontrar trabajos donde se ignora la correlación entre variables de salida.

Por ejemplo, Ugrasen et al. (2014) presentan una investigación sobre la optimización y el análisis de efectos de los parámetros en un maquinado por

descargas eléctricas (EDM) con tres variables de salida del proceso, mediante una red de Retro-Propagación (BP). El nivel de importancia de los parámetros en las salidas es determinado mediante un análisis de varianza (ANOVA), encontrando que la variable con mayor influencia es la  $x_3$ . Los datos de la investigación se presentan en la Tabla 2.1.

**Tabla 2.1 Datos proceso EDM**

RUN	ENTRADAS				SALIDAS		
	$x_1$	$x_2$	$x_3$	$x_4$	$y_1$	$y_2$	$y_3$
1	16	4	3	20	2.3922	5.5224	12
2	16	6	4	25	2.3840	5.5015	11
3	16	8	5	30	2.5260	5.3628	9
4	16	10	6	35	2.7920	7.2172	20
5	20	4	3	30	2.6310	6.8033	19
6	20	6	4	35	2.5410	6.2370	18
7	20	8	5	20	2.8430	7.3504	21
8	20	10	6	25	2.8910	7.4400	23
9	24	4	3	35	3.3170	9.7325	28
10	24	6	4	30	3.2810	8.6313	24
11	24	8	5	25	2.4760	6.0592	13
12	24	10	6	20	2.2530	5.2028	11
13	28	4	3	25	3.0240	8.5863	26
14	28	6	4	20	2.5860	6.3768	17
15	28	8	5	35	2.4820	6.0855	15
16	28	10	6	30	2.3480	5.3757	10

Sin embargo, para obtener conclusiones acertadas mediante el análisis de varianza, se deben cumplir ciertos supuestos sobre los residuales para garantizar que se representa el proceso de manera correcta. Dichos supuestos son:

- Los errores tienen una distribución normal.
- Los errores tienen varianzas iguales.

- Los errores no están auto-correlacionados.

Para verificar el cumplimiento de dichos supuestos, existen diferentes pruebas que comparan un estadístico calculado con cierto valor crítico. Por ejemplo, en el supuesto de normalidad es posible realizar la prueba Anderson-Darling que concluye que los datos son normales si un estadístico calculado  $AD$  es menor a un valor crítico  $VC$ . Para el supuesto de varianzas iguales se verifica mediante la prueba White de homocedasticidad donde se calcula un estadístico  $W$  y se determina que se cumple el supuesto si es menor a un valor crítico. Finalmente, el supuesto de autocorrelación se comprueba mediante la prueba Durbin-Watson donde se busca que el estadístico  $d$  calculado sea mayor al valor crítico.

Mediante dichas pruebas es posible verificar los supuestos en el modelo de Ugrasen et al. (2014), los resultados de estas pruebas se presentan en la Tabla 2.2

**Tabla 2.2 Supuestos con BP**

<b>Residuales</b>	<b>Normalidad</b>	<b>Homocedasticidad</b>	<b>Independencia</b>
$y_1$	$A^2 = 0.129$ $VC = 0.7209$	$W = 0.1173$ $VC = 0.21$	$d = 2.1078$ $VC = 1.7280$
$y_2$	$A^2 = 2.031$ $VC = 0.7209$	$W = 0.1144$ $VC = 0.21$	$d = 0.4276$ $VC = 1.7280$
$y_3$	$A^2 = 0.344$ $VC = 0.7209$	$W = 0.1030$ $VC = 0.21$	$d = 2.3344$ $VC = 1.7280$

En dicha tabla, es posible observar que los residuales de la segunda respuesta no cumplen los supuestos de normalidad e independencia para el modelo BP, por lo tanto, el análisis de varianza no es el indicado para analizar estas respuestas de forma conjunta.

En la Tabla 2.3 se muestran los resultados del análisis de residuales al modelar el proceso con una red de Función de Base Radial (FBR).

**Tabla 2.3 Supuestos con FBR**

<b>Residuales</b>	<b>Normalidad</b>	<b>Homocedasticidad</b>	<b>Independencia</b>
$y_1$	$A^2 = 0.465$ $VC = 0.7209$	$W = 0.1476$ $VC = 0.21$	$d = 2.3012$ $VC = 1.7280$
$y_2$	$A^2 = 0.357$ $VC = 0.7209$	$W = 0.1446$ $VC = 0.21$	$d = 2.4077$ $VC = 1.7280$
$y_3$	$A^2 = 0.186$ $VC = 0.7209$	$W = 0.1326$ $VC = 0.21$	$d = 2.5351$ $VC = 1.7280$

En dicha tabla se verifica el cumplimiento de todos los supuestos para este análisis.

Sin embargo, al tener las mismas variables de entrada para varias variables de salida debe considerarse un análisis multivariado. Para determinar si el análisis debe ser univariado o multivariado es útil realizar la prueba de correlación de Pearson, en la cual una correlación significativa indica que los cambios en una variable de respuesta afectan a otra, y la modelación debe considerarse multivariada.

Sin embargo, dicha condición no suele tomarse en cuenta. Publicaciones como la de Mollah et al. (2008), Ahilan (2013), Jiang et al. (2014), Cevallos (2008) o el mismo Ugrasen et al. (2014) utilizan redes neuronales para modelar procesos de manufactura con varias respuestas considerándolas independientes, y no solo en aplicaciones de redes neuronales, también se ignora la relación entre respuestas en aplicaciones estadísticas como Paulo de Paiva et al. (2014), Lopes et al. (2013), Gomes et al. (2013) y Patel et al. (2016).

Aplicando la prueba de correlación de Pearson en los datos del proceso EDM, se puede observar una correlación fuerte entre las variables ya que los valores son muy cercanos a la unidad. Los resultados se muestran en la Tabla 2.4.

**Tabla 2.4 Correlación proceso EDM**

<b>Correlación Pearson</b>	$y_1$	$y_2$
$y_2$	0.9737	
$y_3$	0.9205	0.9638

Es posible confirmar la correlación significativa por medio de los valores críticos para el coeficiente de correlación de Pearson, para  $n - 2$  grados de libertad y  $\alpha = 0.05$  el  $VC = 0.4260$ , al ser menor este valor crítico, se considera correlación significativa. Además, los valores  $P$  de probabilidad indican también correlación significativa. Dichos valores se muestran en la Tabla 2.5.

**Tabla 2.5 Valores de probabilidad proceso EDM**

<b>P-value</b>	$y_1$	$y_2$
$y_2$	0.0000	
$y_3$	0.0000	0.0000

Como se mencionó anteriormente, mediante el análisis realizado por Ugrasen et al. (2014), por medio de un ANOVA se determinó que la variable con mayor influencia en el proceso es  $x_3$ , sin embargo, no se pueden emitir este tipo de conclusiones. Dados los puntos expuestos anteriormente, se propone el uso del análisis multivariado de varianza (MANOVA) verificando si se cumplen los supuestos para este tipo de análisis, estos son:

- Las respuestas siguen una distribución normal conjunta.
- Las respuestas tienen dependencia entre sí.
- Las respuestas tienen varianzas iguales.

- Las respuestas no están auto-correlacionadas.

Sin embargo, la dificultad se encuentra en las implicaciones de considerar varias respuestas en la red FBR. Por ejemplo, la relación de entradas y salidas de la red está dada por (2.1).

$$y(x) = \sum_{i=1}^n w_i G(\|x - c_i\|) + b + \varepsilon \quad (2.1)$$

Donde  $w$  es el vector de pesos,  $x$  los patrones de entrada,  $c$  los centros o centroides de la red,  $\varepsilon$  es el error,  $y(x)$  es la salida o respuesta del proceso. Escribiendo la ecuación (2.1) en forma matricial:

$$y(x) = \mathbf{G}w \quad (2.2)$$

La ecuación (2.2) genera un vector de salida de la red, pero **¿qué pasará en los pasos siguientes de la red si  $y(x)$  es una matriz?**

Entre la capa de entrada y la capa intermedia se calcula  $\|x - c\|$ , el cual es una distancia  $r$  entre los patrones de entrada y sus respectivos centros.

$$r = \|x - c\| \quad (2.3)$$

Los centros o centroides de la red son un punto fijo dentro de un grupo de datos que representa a dicho grupo. Por lo tanto, esta distancia debe ser mínima para que el centroide represente mejor los datos y así la función de transferencia radial crezca. En otras palabras, cuando  $x - c = 0$  la función devuelve 1, entonces, la distancia entre el vector de centros y el vector de entrada determina la salida de la red.

Existen muchos métodos para determinar estos centros, ya sea métodos de agrupamiento, simulación, aleatorios, redes neuronales, etc. Sin embargo, por estos métodos se genera inestabilidad en la red. Para eliminar esta condición, se propone el algoritmo genético, el cual es un algoritmo de optimización basado en los procesos de evolución biológica (Holland, 1975) y forma parte de los sistemas inteligentes. Dicho algoritmo, evalúa una población generada mediante una función de evaluación, hasta encontrar los centroides de la red. Como función de evaluación se propone maximizar alguna métrica estadística como el coeficiente de determinación  $R^2$  el cual es una métrica de evaluación global (Montgomery et al. 2006).

$$R^2 = 1 - \frac{y'y - w'Gy}{y'y - \frac{(\sum_{i=1}^n y_i)^2}{n}} \quad (2.4)$$

Esto permite al Algoritmo Genético encontrar los centros que mejor aproximen los patrones de entrada a las salidas deseadas maximizando el valor del coeficiente de determinación. Pero al tener varias respuestas **¿cómo se vería afectada la selección de centros?**

Retomando la relación dada en la ecuación (2.3), se puede definir:

$$G(\|x - c\|) = G(r) \quad (2.5)$$

$G$  es una matriz de interpolación dada por:

$$G = \begin{bmatrix} G(x_1 - c_1) & G(x_1 - c_2) & \cdots & G(x_1 - c_m) \\ G(x_2 - c_1) & G(x_2 - c_2) & \cdots & G(x_2 - c_m) \\ \vdots & \vdots & \ddots & \vdots \\ G(x_n - c_1) & G(x_n - c_2) & \cdots & G(x_n - c_m) \end{bmatrix} \quad (2.6)$$

Dicha matriz es determinada por las funciones de base radial, de las cuales existen muchas, las más utilizadas son por ejemplo la Gaussiana:  $G(r) = \exp(r^2)$ , o la tipo Spline:  $G(r) = r^2 \ln r$ .

Rajendra and Shankar (2015), utilizaron una red de base radial para la detección de grietas en la fabricación de vigas voladizas. Diseñaron la red con dos respuestas, localización y profundidad de la grieta, y una función gaussiana de la forma:

$$\varphi(x) = \exp\left(\frac{1}{2\sigma^2}r^2\right)$$

Dicha función también fue utilizada por Mollah y Kumar (2008), en el análisis de un proceso de soldadura TIG, modelando cinco entradas y cuatro salidas.

Praga-Alejo et al. (2013) diseñaron una red de base radial para predicción de un proceso de soldadura GMAW utilizando la función:

$$\varphi(r) = r^2 \ln(r)$$

En base a esto es posible plantear la pregunta ¿cómo seleccionar la función que mejor represente el proceso? En De León-Delgado et al. (2015) se propone utilizar el coeficiente de determinación como función de evaluación en el Algoritmo Genético. Pero **¿funcionarán igual estas funciones en procesos de varias respuestas?, o ¿se deberán usar funciones radiales multivariadas?**

Por su parte, la estimación de los pesos  $w$  es por medio de mínimos cuadrados, ya que este método minimiza la suma de cuadrados del error entre las observaciones  $y$  y los valores  $\hat{y}$  predichos.

$$\hat{w} = (\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'y \quad (2.7)$$

### **¿Este estimador funcionará igual en el caso multivariado?**

Por lo tanto, el objetivo de la presente investigación es desarrollar una metodología general para la predicción en procesos de manufactura, por medio de una red neuronal de Función de Base Radial e incorporando técnicas estadísticas como el análisis de varianza univariado y multivariado, analizando la relación de las variables de salida del modelo y verificando el cumplimiento de los supuestos para el análisis de varianza ya sea univariado o multivariado, con el fin de lograr una mejor predicción del proceso.

Tomando en cuenta el problema descrito se plantean los siguientes aspectos:

#### **2.2 Preguntas de investigación**

- ¿Se puede realizar un análisis de varianza multivariado en la red de Función de Base Radial?
- ¿Qué supuestos debe cumplir la red de base radial para el análisis multivariado de varianza?
- ¿Cómo será la estimación de los centros de la red en el caso multivariado?
- ¿Cómo será la estimación de los pesos de la red en el caso multivariado?
- ¿Qué función de base radial describe mejor el proceso de fundición?

- ¿Qué métrica estadística multivariada es la indicada para validar la predicción?

## **2.3 Hipótesis**

### **➤ Hipótesis 1**

La red neuronal de Función de Base Radial generalizada mediante técnicas estadísticas multivariadas cumplirá con los supuestos para el análisis multivariado de varianza.

### **➤ Hipótesis 2**

Mediante la red neuronal de Función de Base Radial multivariada es posible determinar qué variable de entrada es más importante en el proceso.

## **2.4 Objetivos**

### **2.4.1 Objetivo general**

Desarrollar un modelo generalizado de red neuronal de Función de Base Radial que cumpla con los supuestos para el análisis de varianza multivariado.

### **2.4.2 Objetivos específicos**

- Desarrollar una metodología general para analizar procesos de manufactura de varias respuestas.
- Aplicar metodologías estadísticas multivariantes al modelado y validación de la red FBR.
- Obtener los centros para la función de base radial sin generar inestabilidad en la predicción de la red.
- Mejorar la estimación de los pesos de la red mediante técnicas estadísticas multivariadas.

- Analizar los efectos de diferentes funciones de base radial en la red neuronal multivariada.
- Aplicar la metodología en el proceso de fundición en molde permanente para su validación.

## **2.5 Justificación**

El uso de modelos de predicción tiene su importancia en la optimización de los procesos de manufactura. Sin embargo, esto no sucede en la práctica donde muchas veces se utiliza “prueba y error” para mejorar los procesos, esto es muy común en las empresas por su “sencillez”, pero generan gastos en todos los sentidos: material, costos, tiempo, etc. (Benyounis y Olabi, 2008). Para evitar estos costos es útil hacer uso de herramientas de predicción y así conocer los parámetros indicados del proceso y mantener controladas las características de calidad del producto. Además, dado que en la mayoría de los procesos de manufactura se presentan varias respuestas, y es muy importante encontrar los mejores parámetros para ajustar, predecir y optimizar las respuestas del proceso, se deben tomar en cuenta la relación de todas las variables significativas. Autores como Ahilan et al. (2013), Mollah et al. (2008), Jiang et al. (2014) presentan análisis univariados de diferentes procesos de manufactura sin tomar en cuenta la relación entre sus respuestas para modelar en forma conjunta. El hecho de ignorar tal relación puede traer diversas consecuencias entre las cuales están:

- La falta de ajuste del modelo, si el modelo predice bien para una respuesta es posible que no ajuste adecuadamente para las demás.
- Según las características del proceso, los parámetros de entrada deben ser los mismos para todas las respuestas del modelo, estos parámetros no se encuentran tratando las respuestas de manera independiente.
- Se produce un porcentaje de error en la respuesta real respecto a la estimada.

En general, no se logra la optimización adecuada en las respuestas, o se tomarían decisiones equivocadas en el proceso por no considerar todas las variables de salida. Otra consecuencia es mencionada por Salmasnia et al. (2012) en donde se propone un método que realiza una combinación lineal entre las respuestas, sin embargo, los autores mencionan que este método no encuentra valores óptimos para cada respuesta de manera individual.

La red neuronal de Función de Base Radial es un modelo con gran calidad de predicción por su condición de “aproximador universal” (Haykin 1999, Cevallos 2008), este modelo despliega una simetría radial con respecto a los valores de entrada y así lograr las salidas deseadas. Para esto es necesaria una fase de entrenamiento donde la red determina el valor de los pesos que influyen en las salidas deseadas. Dichas redes trabajan con tres capas, una capa de entrada donde se reciben los datos a tratar, una capa intermedia donde se aplican las funciones de base radial y una capa de salida donde se determinan los resultados de la red. Lo anterior trae consigo diversas ventajas respecto a otras redes neuronales:

- Son buenos predictores en la no linealidad común de los procesos de manufactura.
- Tienen una estructura más simple respecto a otras redes, que le permite una fase de entrenamiento más rápida.
- Con las variables de entrada  $X$  se determinan los centros de la red, gracias a esto la red trabaja en una región de los datos de entrada, que lo convierte en un buen predictor para nuevas observaciones.
- Mediante mínimos cuadrados es posible determinar los pesos de la red, encontrando aquellos que minimizan el error en la salida de la red, reduciendo el tiempo de entrenamiento de otras redes neuronales.

En un proceso de manufactura real es común que exista más de una característica de calidad a ser controlada. En el caso en que no exista correlación entre dichas características, realizar modelos independientes para cada respuesta ayuda a llegar a conclusiones correctas. Sin embargo, si las respuestas están correlacionadas o el uso de modelos independientes no es adecuado, es necesario utilizar un modelo general para representar adecuadamente el proceso.

En otras palabras, cuando hay más de una característica de calidad como variable de salida  $y$  en el proceso se dice que es multivariado. El análisis multivariado es más complejo que uno univariado, porque considera no solo la relación entre las variables predictoras  $X$  con sus variables dependientes  $y$ , sino que también la relación entre múltiples variables dependientes contenidas en una matriz  $Y$ . Las métricas de predicción ayudan a determinar la validez de los modelos aislados cuando se modela para cada respuesta por separado. Pero si no se obtienen resultados adecuados, el uso de modelos aislados no es apto para tomar decisiones acerca de un proceso de manufactura con características de calidad dependientes.

## **2.6 Delimitaciones**

La presente investigación está enfocada en desarrollar una metodología para analizar procesos de manufactura que generen varias características de calidad dependientes, mediante un modelo de red neuronal de función de base radial que tome en cuenta la relación entre varias variables dependientes, mediante el análisis multivariado. Es objetivo encontrar un modelo óptimo para analizar el proceso, sin llegar a optimizar el proceso en sí. La investigación no desarrolla metodologías para el control de los parámetros una vez encontrados.

## **Capítulo 3**

### **Estado del arte**

En el capítulo dos se analizó el trabajo realizado por Ugrasen et al. (2014), en el cual buscaron optimizar los parámetros en un proceso de maquinado EDM, y mediante un ANOVA determinaron la variable más importante. El proceso se trabajó con cuatro variables de entrada y tres de salida. Las salidas presentan una correlación significativa entre todos los pares de variables de respuesta cumpliendo así el supuesto de linealidad entre respuestas para el análisis multivariado de varianza, sin embargo, los autores no mencionan tal relación.

El caso de Ugrasen et al. (2014) no es el único donde se modela un proceso de manufactura mediante una red neuronal y se realiza el análisis de varias respuestas sin considerar analizarlas de forma conjunta.

Ahilan et al. (2013) utilizaron un modelo de red neuronal aplicando los algoritmos de retro propagación, algoritmo genético y optimización por enjambre de partículas, comparando velocidad y precisión. Esto aplicado en la optimización y predicción de un proceso de maquinado CNC. Trabajando con 4 variables de entrada (velocidad de corte, velocidad de alimentación, profundidad de corte y radio de punta) y dos de salida (consumo de energía y rugosidad superficial). Los autores mencionan que el algoritmo que mejor

trabaja con la red neuronal es el algoritmo genético. La influencia de los parámetros del proceso se calculó mediante el análisis de varianza (ANOVA), verificando el supuesto de normalidad para este tipo de análisis por medio de la prueba Anderson-Darling. Sin embargo, no se menciona si se cumplen los supuestos para dicho análisis.

Jiang et al. (2014) utilizan un modelo de red neuronal de retro propagación para investigar la relación entre los parámetros del proceso de curado de polímeros con la viscosidad y el tiempo de curado, y su posterior optimización por medio de un algoritmo genético. La adecuación del modelo se establece por un ANOVA en el cual se miden la interacción de los efectos de los parámetros del proceso en el rendimiento de mezclado. Los autores no mencionan si el modelo cumple los supuestos para este análisis, aun cuando el proceso tiene respuesta bivariada. Por lo tanto, se recomienda verificar el cumplimiento de los supuestos para el análisis multivariado de varianza y analizar las respuestas de manera conjunta.

Los autores Mollah et al. (2008) presentaron dos aplicaciones en procesos de manufactura de la red neuronal de Función de Base Radial en combinación con un algoritmo de Retro-Propagación y un algoritmo genético para comparar los resultados. Se aplicó en los procesos de soldadura TIG y maquinado por flujo abrasivo. El algoritmo genético obtuvo mejores resultados en combinación con la función de base radial. Sin embargo, mencionan que ambos procesos de soldadura presentan respuestas múltiples, por lo tanto, se recomienda el uso de técnicas estadísticas para validar los resultados.

Cevallos (2008) Realizó un estudio comparativo entre las redes neuronales y otras herramientas utilizadas en mejora de la calidad, en base a publicaciones previas de diversos autores. En uno de ellos buscaban una optimización para respuestas múltiples, experimentando con tres variables de

entrada y seis de salida. Cevallos encontró que la red neuronal mejoraba la investigación realizada anteriormente, sin embargo, las respuestas se encontraban significativamente correlacionadas, cumpliendo con el supuesto de linealidad para el análisis multivariado de varianza.

Además de la aplicación de redes neuronales, otros autores utilizan herramientas estadísticas en procesos de manufactura analizando la relación entre respuestas múltiples.

Sampaio et al. (2005) comparan los resultados de una optimización por algoritmo genético con una superficie de respuesta, analizando un proceso de soldadura GMWA. Para esto se tomaron en cuenta tres variables de entrada y cuatro de salida considerándolas independientes, sin embargo, como se mostró en el capítulo dos las respuestas están correlacionadas de una manera significativa, por lo tanto, es importante realizar el análisis tomando en cuenta tal relación.

González et al. (2011) demostraron la importancia de la correlación en un proceso de soldadura GMWA en el que se utilizaron tres variables de entrada y tres de salida. Se analizan los resultados de una regresión múltiple desarrollando un modelo para cada respuesta del proceso y una regresión multivariada tomando en cuenta la correlación entre las respuestas. Se obtienen mejores resultados en la regresión multivariada.

Una preocupación de las industrias de hoy es encontrar aplicaciones eficientes para métodos de optimización basados en modelos. De Paiva et al. (2014) presentaron una metodología de optimización multi-objetivo considerando un análisis de componentes principales aplicado a un proceso de maquinado en CNC considerando dos respuestas de salida.

Lopes et al. (2013) presentan una estrategia para estimar la incertidumbre total que afecta a todas las variables de respuesta, por medio de la inversa de la incertidumbre multivariada como una matriz de ponderación para componentes principales y reemplazar así el conjunto de datos multivariados, con el objetivo de maximizar las métricas  $R_{adj}^2$  y  $R_{pred}^2$ .

Las redes neuronales tienen distintas aplicaciones en predicción y optimización en distintas áreas, por ejemplo:

Rajendra et al. (2015) introducen un nuevo tipo de red para valores complejos usada para la identificación de grietas múltiples en una viga voladiza. Analizando la profundidad y la localización de la grieta. La precisión de la red propuesta fue significativamente mejor comparada con la función de base radial.

Huda et al. (2014), desarrollaron un método híbrido para la detección de señales fuera de control en procesos de manufactura multivariados. Utilizando una red neuronal artificial para aproximar las medidas y luego combinarlo con el aproximado de máxima relevancia. Como resultado, las aproximaciones estimadas fueron más precisas que otras aproximaciones.

Mateo et al. (2013) proponen un método basado en las redes neuronales artificiales de Función de Base Radial para reducir las señales QRST en los electrocardiogramas para el diagnóstico médico. La red se utilizó para proyectar  $x$  en  $y$ , en dos respuestas de salida. Se determinó el coeficiente de correlación entre los valores obtenidos y los reales, para evaluar los modelos. Además, se utilizaron los cuadrados medios del error (MSE) para evaluar el modelo. No se tomó en cuenta la correlación entre las salidas y se evaluó el modelo mediante los MSE simples y no multivariados.

Los autores Lee y Sung (2013), realizaron un estudio sobre la quiebra de empresas coreanas. El estudio sugería un modelo específico de predicción, seleccionando las variables independientes apropiadas. Comparando los resultados de la red neuronal de Retro Propagación y un análisis discriminante multivariado.

Du et al, (2010) desarrollaron una red de base radial para salidas múltiples utilizando un algoritmo recursivo y mínimos cuadrados ortogonales, pero no analiza la relación entre las respuestas.

En el área de la agricultura también se han hecho comparaciones entre métodos tradicionales y redes de base radial para la predicción, en el caso de Cervantes et al. (2012) se buscó predecir la evapotranspiración de referencia encontrando mejores resultados con las redes de base radial que con el método de Priestley y Taylor.

Una vez más se compara la metodología de superficie de respuesta con las redes neuronales artificiales que, en combinación con algoritmos genéticos, obtuvieron mejores resultados que la superficie de respuesta en la predicción del acoplamiento oxidativo de metano sobre Na–W–Mn/SiO<sub>2</sub> a 0.4 MPa. Desarrollando un modelo con tres entradas y dos salidas para optimizar un reactor de lecho fijo isotérmico, tomando ambas salidas como independientes entre sí (Sadeghzadeh et al. 2013).

Las redes neuronales artificiales son una herramienta útil para diversas aplicaciones, por ejemplo, el modelado del proceso de extracción de productos de agricultura (Karimi et al. 2012). Una red neuronal artificial entrenada encontró el punto para minimizar la humedad y maximizar la velocidad de secado, energía y la eficiencia de energía, para obtener los valores deseados. Los resultados indican que al obtener valores altos de  $R^2$  y  $R_{adj}^2$ , y al mismo

tiempo valores bajos para una superficie de respuesta, se representa una alta dependencia entre los valores de predicción.

Se han utilizado funciones de deseabilidad y algoritmos genéticos para la optimización de un problema de múltiples respuestas correlacionadas. Salmasnia et al. (2012).

Liu et al. (2014) presentan un algoritmo de optimización global mediante un esquema de rectángulos divididos (DIRECT). Comparando los resultados de una red de base radial y una superficie de respuesta, la metodología DIRECT resultó una estrategia que acelera la optimización en términos de eficiencia y precisión. Aunque los tres algoritmos propuestos obtuvieron un buen desempeño, el esquema DIRECT fue más amigable con los cálculos computacionales llegando más rápido al objetivo.

Jie et al. (2014) presentan un metamodelo basado en optimización global (AMGO) adaptable. Este algoritmo es un modelo híbrido compuesto por un Kriging y una función de base radial. Se aplicó el algoritmo para aproximación en problemas de tipo caja negra reduciendo los efectos desfavorables de utilizar algún metamodelo de manera incorrecta. Los resultados indican que el método propuesto resultó satisfactorio en la precisión y eficiencia de la convergencia del proceso iterativo con un bajo costo computacional.

En el trabajo publicado por Shahlaei et al. (2015) se describe una aproximación simple del monitoreo del celecoxib en sueros humanos, a través de un análisis por componentes principales combinado con una red de Función de Base Radial. Los resultados indican que el método propuesto es una buena alternativa para el monitoreo de este antiinflamatorio.

Haji et al. (2015) proponen un esquema simple de optimización para resolver problemas llamados de descomposición, para cuantificar las correlaciones de variables en problemas de optimización compatibles con varios tipos de problemas. Comparando los casos de optimización directa sin descomposición con diferentes niveles del mismo. Esto porque los métodos desarrollados mediante funciones de base radial de alta dimensión proporcionan información sobre variables correlacionadas, aplicables a problemas que pueden ser completamente descompuestos. Los resultados indican que los problemas con variables correlacionadas fuertemente no se pueden descomponer.

En el trabajo de Tyan et al. (2014) se presenta un modelado global para fidelidad de variables que amplía el modelo de variables complejas. Utilizando un diseño de experimentos para probar las funciones de alta y baja fidelidad e iniciar la función de escala utilizando una función de base radial. De manera que se hace posible eliminar la evaluación de gradiente de alta fidelidad del proceso, haciéndolo más eficiente para problemas de diseño de alta dimensión.

Akhtar et al. (2015) proponen un algoritmo basado en superficie de respuesta para una aproximación multi-objetivo aplicable a problemas de caja negra. Utilizando una función de base radial para calcular iterativamente aproximaciones de la función objetivo. Los resultados indican que entre mayor número de iteraciones tenga, el método propuesto mejorará la aproximación.

El objetivo del trabajo de Gomes et al. (2015) es presentar una aplicación de una red neuronal de Función de Base Radial gaussiana para el sistema de diagnóstico para la identificación de problemas en plantas nucleares con reactor de agua a presión. Concluyendo que, de acuerdo a la interpolación por la

función de base radial, la red tiene un desempeño satisfactorio para aplicarse a la predicción de accidentes nucleares.

### **3.1 Conclusiones del estado del arte**

Las tecnologías inteligentes han sido objeto de diversas aplicaciones en años recientes incrementando su interés en la comunidad científica. Estos sistemas inteligentes están basados en el comportamiento humano y fueron desarrollados para resolver problemas y tomar decisiones en problemas complejos (Yen y Langari 2000). Uno de estos sistemas inteligentes son las redes neuronales, las cuales se desarrollaron como una forma de modelar el proceso biológico de aprendizaje del cerebro humano para resolver problemas (Qasem y Shamsuddin 2011). Tienen aplicaciones en ciencia e ingeniería principalmente en interpolación, reconocimiento de patrones y clasificación (Rajendra y Shankar 2015) ya que son una herramienta útil para establecer relaciones entre variables, analizar datos complejos o con información incompleta (Shahlaei et al. 2015). Existen diferentes tipos, la más utilizada es el gradiente descendente basado en Retro-Propagación, sin embargo, tiene la desventaja de una convergencia lenta y su facilidad de estancarse en un mínimo local (He et al. 2009, Rajendra & Shankar 2015).

Además de las redes de Retro-Propagación otros “aproximadores universales” son las redes de Función de Base Radial, las cuales utilizan este tipo de funciones para su activación y son una herramienta muy utilizada en la industria de manufactura debido a su capacidad de representar correctamente el proceso. Para lograrlo es muy importante determinar cuidadosamente los centros de la función de base radial. Como su nombre lo indica, utiliza funciones de base radial para su activación, y los pesos de la red pueden estimarse por cálculos lineales. Esto permite basarse en características locales del espacio de entrada para un aprendizaje más rápido (He et al. 2009) reduciendo los costos de computación requeridos por otras metodologías.

Otros sistemas inteligentes son los Algoritmos Genéticos, los cuales pueden combinarse con las redes de Función de Base Radial (He et al. 2009, Jiang et al. 2014, Gomes y Canedo 2015), principalmente para optimizar los centros de la función de base radial, y de esta forma disminuir la aleatoriedad causada por la determinación de dichos centros mediante otras herramientas.

La aplicación principal de las redes FBR es la predicción en diferentes casos, diversos autores han determinado modelos que representan de buena forma su problemática. Estas problemáticas requieren analizar una o varias respuestas, pero la mayoría de los autores no mencionan si existen más variables en los casos donde se toma en cuenta solo una respuesta. Incluso en problemas de varias respuestas se utiliza un ANOVA para determinar la adecuación del modelo y no se menciona si este método es el adecuado o pudo utilizarse el análisis de varianza multivariado.

Los procesos de manufactura generalmente presentan más de una característica de respuesta, y para modelar un proceso con varias respuestas generalmente se utilizan modelos independientes para cada respuesta. Si la correlación entre respuestas no es significativa, analizarlas de manera independiente puede resultar útil. Pero si tales respuestas tienen correlación, la forma de llegar a las mejores conclusiones será analizar las variables de manera conjunta.

Se busca revisar los casos del análisis de varias respuestas, si los modelos cumplen los supuestos para el análisis de varianza multivariado, y en caso de cumplirlos aplicarlo junto con herramientas de la estadística multivariada.

## **Capítulo 4**

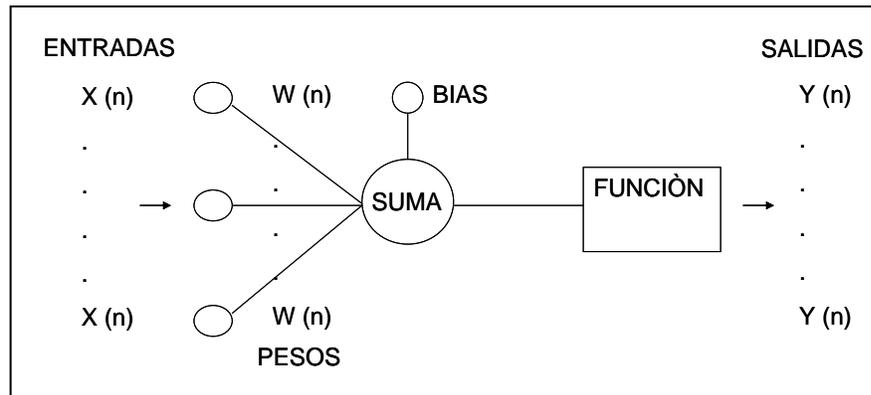
### **Marco teórico**

#### **4.1 La red neuronal artificial**

Las redes neuronales artificiales forman parte de los sistemas inteligentes y su principal característica es que basan su funcionamiento en los sistemas biológicos, el comportamiento humano y la naturaleza en general. Además de las redes neuronales existen los algoritmos genéticos (evolución) y la lógica difusa (razonamiento humano). Por su parte, las redes neuronales están basadas en el comportamiento de las neuronas en el cerebro humano.

Una neurona es una unidad de procesamiento de información fundamental en la red neuronal. Un modelo de neurona identifica tres elementos básicos (Haykin, 1999):

1. La sinapsis.
2. Los pesos.
3. Una función de activación.



**Figura 4.1 Estructura de una red neuronal**

Otra característica importante, y que puede observarse en la estructura de la red mostrada en la Figura 4.1 es el bias o umbral, el cual depende del efecto de la función de activación, incrementando o disminuyendo si el umbral es positivo o negativo respectivamente. Un modelo de red neuronal se representa matemáticamente por:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (4.1)$$

y

$$y_k = \varphi(u_k + b_k) \quad (4.2)$$

Donde  $x_1, x_2, \dots, x_m$  son los patrones de entrada;  $w_{k1}, w_{k2}, \dots, w_{km}$  son los pesos sinápticos de la neurona  $k$ ;  $u_k$  es la combinación lineal de salidas debido a las señales de entrada;  $b_k$  es el bias;  $\varphi(\cdot)$  es la función de activación; y  $y_k$  es la señal de salida de la neurona.

## 4.2 Funciones de activación

Una neurona propaga un valor de salida a través de conexiones unidireccionales hacia otras células de la red. Matemáticamente se utiliza una función de base para dar un valor de activación. Existen tres tipos básicos de funciones de activación (Haykin, 1999):

1. Función umbral: esta función explica que durante un instante cuantizado de tiempo la neurona responde a la actividad de su sinapsis, verificando si dicha sinapsis alcanza o excede un umbral fijo, si lo hace la neurona se activa (McCulloch & Pitts, 1943).
2. Función lineal: establece la región de  $v$  entre cero y uno.
3. Función sigmoideal: se define como una función creciente con un equilibrio entre comportamiento lineal y no lineal.

También existen funciones de base, donde una de las capas de la red utiliza un conjunto a ser aproximada por la función. Un ejemplo es la Función de Base Radial.

## 4.3 Funciones de Base Radial

Las funciones de base radial se caracterizan por su respuesta que aumenta o disminuye con respecto a la distancia entre los patrones de entrada a un punto fijo llamado centro o centroide mediante una función de simetría radial (Broomhead & Lowe, 1988).

Es una función de segundo orden o de tipo híper-esférico, utiliza en la capa oculta funciones radiales no lineales con sus centros gravitacionales propios, mientras que la capa de salida se compone por funciones lineales. Se le conoce como modelo híbrido ya que utiliza aprendizaje supervisado y no supervisado.

#### 4.4 Interpolación por una red de Función de Base Radial

Una característica importante de la red de Función de Base Radial (FBR) es su buen rendimiento en tareas con patrones complejos, transformando dichos patrones de un espacio de alta dimensión a una forma no lineal, utilizando el teorema de separabilidad de patrones de Cover (1965).

Dicho teorema supone que un patrón  $x$  es un vector en un espacio de entrada  $m_0$ -dimensional, entonces un vector  $\varphi(x)$  proyecta puntos del espacio de entrada a sus correspondientes puntos en un nuevo espacio de dimensión  $m_1$ .

El teorema de Cover presenta dos enfoques básicos:

1. La formulación no lineal de la función oculta definida por  $\varphi_i(x)$ , donde  $i = 1, 2, \dots, m_1$ , y  $x$  es el vector de entrada.
2. Un espacio oculto de dimensión más alta comparada con el espacio de entrada, y determinada por el valor asignado a  $m_1$ .

En algunos casos el uso de una proyección no lineal puede ser suficiente para producir separabilidad lineal sin incrementar la dimensionalidad o el espacio de la unidad oculta.

Considerando una red de propagación hacia adelante con tres capas eligiendo una unidad de salida única para simplificar, la red se diseña para realizar una proyección no lineal en el espacio de salida de la neurona oculta, seguido de una proyección lineal del espacio oculto al espacio de salida. De manera general la red representa una proyección del espacio de entrada  $m_0$ -dimensional hacia el espacio de salida unidimensional. Representado como:

$$s: \mathbb{R}^{m_0} \rightarrow \mathbb{R}^1$$

La fase de entrenamiento y la fase de generalización del proceso de aprendizaje de la red se definen como (Broomhead & Löwe, 1988):

- La fase de entrenamiento constituye la optimización de un proceso de ajuste para la superficie  $\Gamma$ , basado en los puntos conocidos presentados a la red en forma de ejemplos de entradas-salidas (patrones).
- La fase de generalización es sinónimo de interpolación entre los puntos de datos, la interpolación se realiza a lo largo de la superficie limitada generada por el proceso de ajuste como una aproximación óptima hacia la verdadera superficie  $\Gamma$ .

Esto lleva a la teoría de interpolación multivariable en un espacio de alta dimensión. El problema de interpolación está definido por:

Sea  $\{x_i \in \mathbb{R}^{m_0} | i = 1, 2, \dots, N\}$  un conjunto de  $N$  puntos diferentes, y  $\{d_i \in \mathbb{R}^1 | i = 1, 2, \dots, N\}$  sus  $N$  números reales correspondientes, encontrar una función  $F: \mathbb{R}^N \rightarrow \mathbb{R}^1$  que satisfaga la condición de interpolación:

$$F(x_i) = d_i, \quad i = 1, 2, \dots, N \quad (4.3)$$

La superficie de interpolación está restringida a pasar por todos los puntos de entrenamiento.

La técnica por funciones de base radial consiste en elegir una función  $F$  de la siguiente forma (Powell, 1987):

$$F(x) = \sum_{i=1}^N w_i \varphi(\|x - x_i\|) \quad (4.4)$$

Donde  $\{\varphi(\|x - x_i\|) | i = 1, 2, \dots, N\}$  es un conjunto de  $N$  funciones arbitrarias (generalmente no lineales), conocidas como *funciones de base radial*, y  $\|\cdot\|$  denota una distancia o norma. Los puntos conocidos  $x_i \in \mathbb{R}^{m_0}$ ,  $i = 1, 2, \dots, N$  son los centros de la función de base radial.

Introduciendo las condiciones de interpolación de la ecuación (4.3) en la ecuación (4.4), forman unas ecuaciones lineales simultáneas para los coeficientes desconocidos  $\{w_i\}$  (pesos).

$$\begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1N} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{N1} & \varphi_{N2} & \cdots & \varphi_{NN} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \quad (4.5)$$

Donde

$$\varphi_{ji} = \varphi(x_j - x_i), \quad (j, i) = 1, 2, \dots, N \quad (4.6)$$

Los vectores de  $N \times 1$   $F(x)$  y  $w$  representan la respuesta y el vector de pesos lineales respectivamente,  $N$  es el tamaño de la muestra de entrenamiento. La matriz de  $N \times N$  denota  $\Phi$  cuyos elementos  $\varphi_{ji}$  son:

$$\Phi = \{\varphi_{ji} | (j, i) = 1, 2, \dots, N\} \quad (4.7)$$

Esto es la matriz de interpolación. Reescribiendo la ecuación (4.5):

$$\Phi w = x \quad (4.8)$$

Asumiendo que  $\Phi$  es no singular y que su inversa existe, puede resolverse la ecuación (4.8) para  $w$  de la forma:

$$w = \Phi^{-1}x \quad (4.9)$$

Para asegurar que la matriz  $\Phi$  es no singular, se utiliza el teorema de Micchelli.

#### 4.5 Teorema de Micchelli

Se plantea este teorema con un conjunto de puntos diferentes es un espacio  $\mathbb{R}^{m_0}$ , donde la matriz de interpolación  $\Phi$  es no singular. Existen muchas funciones de base radial que aplican el teorema de Micchelli (1986), las más utilizadas y con mayor importancia dentro de las redes neuronales son:

1. Función multi-cuadrática:

$$\varphi(r) = \sqrt{1 + r^2} \quad (4.10)$$

2. Función multi-cuadrática inversa:

$$\varphi(r) = \frac{1}{\sqrt{1 + r^2}} \quad (4.11)$$

3. Función Gaussiana:

$$\varphi(r) = \exp\left(-\frac{r^2}{\sigma^2}\right) \quad (4.12)$$

Para que las funciones anteriores sean no singulares, los puntos  $\{x_i\}_{i=1}^N$  deben ser diferentes, sin importar el tamaño de  $N$  o la dimensionalidad  $m_0$  de los vectores  $x_i$ .

Una característica importante de las funciones inversas multi-cuadráticas y las Gaussianas es que son funciones localizadas, esto en el sentido que  $\varphi(r) \rightarrow 0$  si  $r \rightarrow \infty$ . En otras palabras, si la diferencia entre los puntos  $x_i$  y su centro es muy grande, la salida de la función se aproximará a cero, así la matriz de interpolación  $\Phi$  será positiva definida. En contraste, las ecuaciones multicuadráticas son no localizadas,  $\varphi(r)$  se convierte en ilimitada si  $r \rightarrow \infty$ , y la correspondiente matriz de interpolación  $\Phi$  tiene  $N - 1$  eigenvalores negativos y solo uno positivo, con este resultado la matriz no es definida positiva (Micchelli, 1986). Sin embargo, su no singularidad la vuelve adecuada para su uso en el diseño de la red.

Otras funciones utilizadas en la red son:

Función radial cúbica:

$$\varphi(r) = r^3 \quad (4.13)$$

Función Spline de capa delgada:

$$\varphi(r) = r^2 \log(r) \quad (4.14)$$

Funciones Spline poli-armónicas:

$$\varphi(r) = r^k; \quad k = 5, 7, 9, \dots \quad (4.15)$$

$$\varphi(r) = r^k \log(r); \quad k = 4, 6, 8, \dots \quad (4.16)$$

#### 4.6 Funciones Gaussianas Multivariadas

Si  $G(x, x_i)$ , es una función que satisface la condición:  $G(x, x_i) = G(\|x - x_i\|)$ , entonces se dice que dicha función es linealmente diferenciable. Un ejemplo de esta condición es la función Gaussiana multivariada definida por:

$$G(x, x_i) = \exp\left(-\frac{1}{2\sigma_i^2} \|x - c_i\|^2\right) \quad (4.17)$$

Donde  $c_i$  define los centros de la función y  $\sigma_i$  denota la amplitud.

#### 4.7 Redes FBR generalizadas

La red de Función de Base Radial (FBR), en su forma básica incluye tres diferentes capas (Broomhead & Lowe, 1988):

- Capa de entrada: está formada por los nodos fuente  $x_i$  (unidades sensoriales). El número de nodos está definido por la dimensión  $m_0$  del vector  $x$ .
- Capa intermedia: es una capa oculta de gran dimensión y en la que las unidades (neuronas) que la forman son las funciones base para los datos de entrada  $\varphi_i(x_i)$ .
- Capa de salida:  $d_j$  que tiene la responsabilidad en la red de la activación de patrones aplicados en la capa de entrada.

La estructura básica de la red es mostrada en la Figura 4.2.

Generalmente, el número de funciones base es menor que el número de datos ( $m_1 \leq N$ ). Pensando en las funciones de base radial, se define:

$$\varphi_i(x) = G(\|x - c_i\|), \quad i = 1, 2, \dots, m_1 \quad (4.18)$$

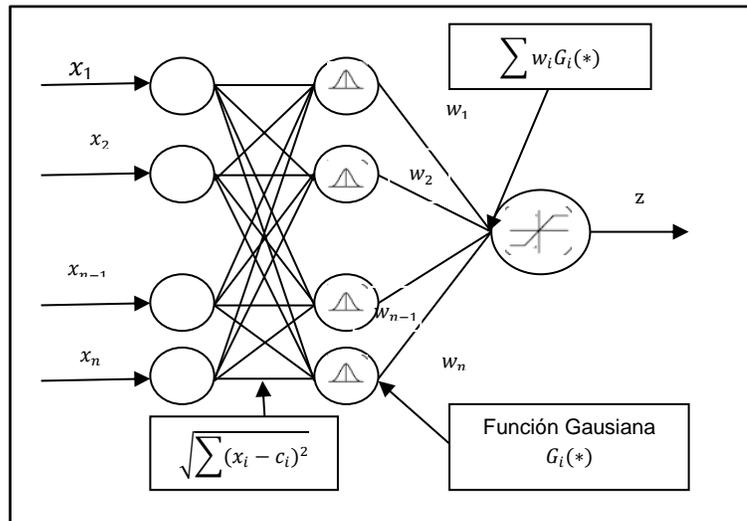


Figura 4.2 Estructura de la red FBR

Donde los centros  $\{c_i | i = 1, 2, \dots, m_1\}$  son determinados. La función de base que se elegirá, es aquella en que se pueda explicar la relación *entradas-salidas* de la red correctamente de la forma:

$$y(x) = \sum_{i=1}^n w_i G(\|x - c_i\|) + b + \varepsilon \quad (4.19)$$

$\|x - c_i\|$  es una distancia  $r$  y dejando a  $G$  definida en base a esta distancia:

$$\mathbf{G} = \begin{bmatrix} G(x_1 - c_1) & G(x_1 - c_2) & \cdots & G(x_1 - c_m) \\ G(x_2 - c_1) & G(x_2 - c_2) & \cdots & G(x_2 - c_m) \\ \vdots & \vdots & \ddots & \vdots \\ G(x_n - c_1) & G(x_n - c_2) & \cdots & G(x_n - c_m) \end{bmatrix} \quad (4.20)$$

#### 4.8 Estimación de los pesos de la red por mínimos cuadrados

La estimación de  $w$  por mínimos cuadrados, minimiza la suma de cuadrados de las desviaciones de las observaciones  $y$  y los valores  $\hat{y}$  predichos. La función de mínimos cuadrados es:

$$\begin{aligned}
SSE &= \sum_{i=1}^n \varepsilon^2_i \\
&= \sum_{i=1}^n \left( y_i - w_b - \sum_{j=1}^q w_j G_{ij} \right)^2
\end{aligned} \tag{4.21}$$

Expresado en notación matricial:

$$\mathbf{y} = \mathbf{G}\mathbf{w} + \boldsymbol{\varepsilon} \tag{4.22}$$

Donde  $y$  es el vector de salida  $n * 1$  resultante de las observaciones,  $G$  es una matriz de  $n * q$  generada por la matriz de centros  $C$  y la matriz  $X$  de observaciones,  $w$  es un vector de  $q * 1$  de los pesos de la red donde  $w_b$  es la estimación del bias, finalmente  $\varepsilon$  es un vector de  $n * 1$  de errores aleatorios. Por lo tanto, se desea determinar el vector  $w$  que minimice la ecuación (4.23):

$$SSE = \sum_{i=1}^n \varepsilon^2_i = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{G}\mathbf{w})'(\mathbf{y} - \mathbf{G}\mathbf{w}) \tag{4.23}$$

Un estimador de  $w$  que minimiza la SSE en (4.23) es la ecuación (4.24):

$$\hat{\mathbf{w}} = (\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{y} \tag{4.24}$$

En (4.24) se asume que  $G'G$  es no singular. Esto generalmente se cumple si  $n > q + 1$  y  $x_j$  no es combinación lineal de otras  $x$ .

En el caso multivariado, el vector de pesos  $\hat{w}$  se convierte en una matriz de pesos  $\hat{W}$ . Esta matriz de pesos es un estimador por mínimos cuadrados de  $W$  ya que minimiza la matriz de suma de cuadrados del error  $E$ :

$$E = \hat{\Xi}'\Xi = (Y - GW)'(Y - GW) \quad (4.25)$$

El estimador por mínimos cuadrados de  $\widehat{W}$  también minimiza  $tr(Y - GW)'(Y - GW)$  y  $|(Y - GW)'(Y - GW)|$  las cuales son cantidades escalares. La traza de la matriz  $E$  se puede expresar como sigue:

$$tr(Y - GW)'(Y - GW) = \sum_{i=1}^n \sum_{j=1}^p \hat{\xi}_{ij}^2 \quad (4.26)$$

#### 4.9 Estimación de los centros de la red por Algoritmo Genético

Debido a que la respuesta de la red de Función de Base Radial depende de una distancia  $r$  entre los datos  $x$  y los centros  $c_i$ , es posible determinar dichos centros por medio de la minimización de esta distancia. El Algoritmo Genético es un método de optimización basado en los procesos de evolución biológica (Holland, 1975), que forma parte de los sistemas inteligentes. Funcionan generando una población aleatoria de individuos, los cuales son potenciales soluciones al problema. Estos individuos cambian probabilísticamente en el tiempo por medio de operadores de cruce y mutación. Un proceso de selección determina qué individuos (entre padres y descendientes) permanecen en la próxima generación.

La mayoría de los AG se componen de:

- Esquema de decodificación. Este esquema permita transformar los puntos de un espacio de parámetros, para representarlos en código. Los cuales forman los cromosomas. (Nelles, 2001).
- Evaluación por aptitud o "fitness". Por medio de una función objetivo se evalúa cada miembro de la población.

- Selección. Esta operación determina cuales padres participan en producir los siguientes hijos para la próxima generación.
- Cruce. Por medio de la combinación de dos individuos aleatorios, se genera un hijo con los rasgos de ambos padres.
- Mutación. Este operador afecta a un individuo en alguno de sus cromosomas, por ejemplo, si se emplea el código binario la mutación afectaría un bit del individuo.

Y se desarrollan de la siguiente manera:

1. Generar una población aleatoria.
2. Evaluar la población.
3. Seleccionar los mejores individuos.
4. Realizar el operador de cruce para encontrar nuevas soluciones.
5. Determinar si algún individuo muta.
6. Generar una nueva población.
7. Regresar al punto dos hasta que se cumpla el criterio de paro.
8. Encontrar una solución óptima o la mejor solución.

Para realizar el paso 2 del algoritmo, se utiliza una función de evaluación para optimizar la población seleccionada. Dicha función determina el objetivo del algoritmo, ya sea maximización o minimización. Como se mencionó anteriormente, la FBR depende de una distancia entre los patrones de entrada y sus centros. Una manera de comprobar esto es utilizando el coeficiente de determinación  $R_q^2$  (Rencher, 2002), el cual está definido por (4.27):

$$R_q^2 = (Y'Y - n\bar{y}\bar{y}')^{-1}(\widehat{W}'G'Y - n\bar{y}\bar{y}') \quad (4.27)$$

Cuando  $q = 1$   $R_q^2$  está en forma escalar por ser el caso univariado. Cuando  $q > 1$   $R_q^2$  se convierte a la extensión directa multivariada del caso univariado. Para convertir  $R_q^2$  a su forma escalar cuando  $q > 1$ , se utiliza la traza  $tr(R_q^2)$  dividida por el número  $q$  de respuestas o columnas de  $Y$ , tal que  $0 \leq tr(R_q^2) \leq 1$ . También Se puede utilizar el determinante  $|R_q^2|$  (Rencher, 2002). Para determinar que la calidad del modelo es buena  $R_q^2$  debe aproximarse a la unidad.

#### 4.10 Distancia de Mahalanobis

La maximización de  $R_q^2$  depende de la distancia  $r$  entre los patrones de entrada y los centros correspondientes. Existen diferentes tipos de distancias, la distancia Euclideana que es la más utilizada en las redes FBR, sin embargo, en el presente trabajo se propone el uso de la distancia de Mahalanobis ( $d_m$ ) definida por:

$$d_m = \sqrt{(x - \mu)' \Sigma^{-1} (x - \mu)} \quad (4.28)$$

Esta distancia utiliza la matriz de varianzas y covarianzas dada por  $\Sigma$  siendo una matriz cuadrada simétrica de tamaño  $p \times p$ . Esta distancia se distribuye como una  $\chi^2$  con  $p$  grados de libertad.

#### 4.11 Análisis Multivariado de Varianza (MANOVA)

El objetivo de la red neuronal es intentar crear una predicción del proceso de manufactura, relacionando una o varias variables de salida  $y$  con una o varias variables de entrada  $x$ . En otras palabras, cada  $y_1, y_2, \dots, y_p$  es predicha por toda  $x_1, x_2, \dots, x_q$ . Con esto surgen dos preguntas importantes: ¿qué tan bien estará explicada la variabilidad relacionada con el proceso?, y ¿cómo saber que las variables incluidas en la red tienen influencia significativa en las respuestas?

La estimación de los pesos de la red por mínimos cuadrados permite hacer inferencia estadística para resolver las preguntas planteadas, realizar pruebas de hipótesis, y en general tomar decisiones del proceso en base a la información obtenida en el modelo. La inferencia estadística requiere que se cumplan ciertos supuestos, que para el caso multivariado se presentan a continuación:

- **Normalidad multivariada**

Existen muchas pruebas para evaluar la normalidad multivariante en la literatura estadística (Mecklin y Mundfrom 2004). Sin embargo, no se conoce una prueba definitiva y se recomiendan varias pruebas para evaluarla.

Henze y Zirkler (1990) introducen una clase de pruebas para determinar la normalidad multivariada. La prueba Henze-Zirkler mide la distancia entre dos funciones de distribución. El estadístico utilizado es:

$$T_{\beta} = n(4I_{(S \text{ singular})} + D_{n,\beta}I_{(S \text{ no singular})}) \quad (4.29)$$

Dónde:

$$D_{n,\beta} = \int \left| \psi_n(\mathbf{t}) - \exp\left\{-\frac{1}{2}\|\mathbf{t}\|^2\right\} \right|^2 \phi_{\beta}(\mathbf{t}) d\mathbf{t} \quad (4.30)$$

$I$  es una función indicadora y  $\psi_n(\mathbf{t})$  es una función característica empírica de las  $\mathbf{Z}_j^*$ s, dadas respectivamente por:

$$\psi_n(\mathbf{t}) = \frac{1}{n} \sum_{i=1}^n \exp\{i\mathbf{t}'\mathbf{Z}_j^*\}, \quad \mathbf{t} \in \mathbb{R}^p \quad (4.31)$$

Y

$$\phi_{\beta}(\mathbf{t}) = \frac{1}{(2\pi\beta^2)^{p/2}} \exp\left\{-\frac{1}{2\beta^2} \|\mathbf{t}\|^2\right\}, \quad \mathbf{t} \in \mathbb{R}^p \quad (4.32)$$

Se rechaza  $H_0$  para valores grandes de  $T_{\beta}$ .

Otra prueba que ha demostrado ser confiable y estable para asegurar normalidad multivariada (Romeu y Ozturk, 1993) es la prueba de simetría y curtosis multivariada de Mardia (Mardia, 1970). El cual estima la simetría multivariada de la forma:

$$\hat{\beta}_{1,p} = \sum_{i=1}^n \sum_{j=1}^n \frac{[(y_i - \bar{y})' S^{-1} (y_j - \bar{y})]^3}{n^2} \quad (4.33)$$

Y la curtosis multivariada:

$$\hat{\beta}_{2,p} = \sum_{i=1}^n \frac{[(y_i - \bar{y})' S^{-1} (y_i - \bar{y})]^2}{n} \quad (4.34)$$

Los cuales convergen en una distribución  $N(\mu, \sigma^2)$  con media  $\mu = p(p + 2)$  y varianza  $\sigma^2 = 8p(p + 2)/n$ . Sustrayendo  $\mu$  de  $\hat{\beta}_{2,p}$  y dividiéndolo por  $\sigma$ , entonces:

$$Z = \frac{(\hat{\beta}_{2,p} - \mu)}{\sigma} \xrightarrow{d} N(0,1) \quad (4.35)$$

El rechazo de la normalidad mediante pruebas de Mardia indica la presencia de puntos atípicos, o que la distribución es significativamente diferente de una distribución normal multivariada. Si no se rechaza, la

distribución se supone normal multivariada. Romeu y Ozturk proponen una corrección de la prueba de Mardia para muestras pequeñas (menos de 50 observaciones) mediante valores críticos empíricos para las pruebas de asimetría y la curtosis.

Por último, se propone la prueba de Shapiro-Wilk. Dicha prueba surge del concepto de invariancia ya que plantea que el vector aleatorio es normal multivariante si y sólo si su estandarización lo es. Los datos estandarizados son:

$$Z_i = S^{1/2}(X_i - \bar{X}) \quad (4.36)$$

Donde  $\bar{X}$  es la media muestral y  $S$  la matriz de covarianzas muestral. El estadístico Shapiro-Wilk multivariante se define como:

$$W_M = \frac{1}{k} \sum_{j=1}^k W_j \quad (4.37)$$

Donde  $W_1, \dots, W_k$  son estadísticos de Shapiro-Wilk de cada componente  $Z_1, \dots, Z_n$ . Se rechazará la normalidad cuando  $W_M$  sea pequeño en comparación a los valores críticos de tablas.

En caso de rechazar las pruebas de normalidad multivariada, hay que identificar valores atípicos o realizar una transformación de los datos para lograr la normalidad multivariante. Un ejemplo puede ser una extensión de la transformación de potencia de Box-Cox propuesta por Andrews et al., (1971), aunque determinar una transformación apropiada es complicado (Bilodeau y Brenner, 1999).

- **Linealidad**

El segundo afirma que el modelo está correlacionado y no necesita vectores  $x$  adicionales para predecir las respuestas. Para probar este supuesto se utiliza el coeficiente de correlación de Pearson mide el grado de relación entre dos variables aleatorias  $Y_i$  y  $Y_j$ , está dado por:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j} = \frac{cov(Y_i, Y_j)}{\sqrt{var(Y_i)var(Y_j)}} \quad (4.38)$$

Donde  $\rho_{ij}$  es una matriz de correlaciones para un vector aleatorio  $Y$  de tamaño  $p$  y  $-1 \leq \rho_{ij} \leq 1$ .

Debido a que la matriz de correlaciones no depende de la escala de las variables aleatorias, es utilizada para expresar relación entre dichas variables medida en diferentes escalas. Si los elementos de  $Y$  son dependientes, su valor de será cero, y si son independientes su valor será uno. Así, el determinante de la matriz de correlación puede interpretarse como una medida de asociación o no asociación (Timm, 2002).

Dado que los procesos de manufactura tienen comúnmente un comportamiento no lineal (Cevallos, 2008), existe la posibilidad de que la relación entre las respuestas sea no lineal. El coeficiente de correlación de Pearson mide la correlación lineal entre las variables. Una prueba menos sensible a la correlación lineal es el coeficiente de correlación  $\tau$  de Kendall. Dicho coeficiente de correlación entre dos variables  $i$  y  $j$  se representa mediante la expresión:

$$\tau_{i,j} = 4 \int_0^1 \int_0^1 C_{i,j} d_i d_j \quad (4.39)$$

Dónde:

$$C_{i,j} = C(1, \dots, 1, F_i(t), 1, \dots, 1, F_j(t), 1, \dots, 1) \quad (4.40)$$

- **Homogeneidad de varianzas**

Si la varianza no es constante invalida las pruebas de significancia estadística, causando así la determinación de estimadores ineficientes.

White (1980) derivó la estimación de una matriz de varianzas y covarianzas consistente que provee una prueba de contraste de los coeficientes con presencia de un patrón no conocido de heteroscedasticidad, la matriz de Varianza Covarianza de White está dada por:

$$\hat{\Sigma}_W = \frac{N}{N-k} (X'X)^{-1} \left( \sum_{i=1}^N \hat{u}_i^2 x_i x_i^T \right) (X'X)^{-1} \quad (4.41)$$

En este caso  $X$  es calculado por la ecuación (4.20). El estimador de homocedasticidad de mínimos cuadrados ordinarios es:

$$V = \sigma^2 [X'X]^{-1} \quad (4.42)$$

Si no hay heteroscedasticidad (4.42) será un estimador consistente de la matriz de varianza y covarianza de los parámetros ( $Var(b)$ ), si existe no lo será. Partiendo de esto White construyó una prueba de contraste a través de obtener el coeficiente de determinación ( $R^2$ ) de una regresión auxiliar de los errores contra las variables explicativas del modelo original, sus cuadrados y sus productos cruzados, luego se construye un estadístico que se distribuye, como una Chi cuadrada con  $p - 1$  grados de libertad ( $NR^2 \sim \chi_{p-1}^2$ ). Donde  $p$  es el

número de regresores de la regresión auxiliar sin incluir la constante. El estadístico provee una prueba de hipótesis de que todas las pendientes de la regresión auxiliar son cero, es decir;

$$H_0: \theta_1 = \theta_2 = \theta_3 = \dots = \theta_p = 0 \quad (4.43)$$

A su vez White señala que es una prueba general de especificación del modelo, dado que la hipótesis nula se fundamenta en que los errores son homoscedásticos e independientes de los regresores y que la especificación lineal es la correcta. Si la hipótesis nula se rechaza, una o más de estas condiciones violadas conducen a una prueba significativa. De lo contrario, la ausencia de significancia estadística de la prueba, puede implicar que ninguna de las tres condiciones es violada.

- **Independencia entre respuestas**

La tendencia a tener residuales positivos y negativos indica una correlación serial positiva. Esto es un indicio de la violación del supuesto de independencia. La aleatorización del experimento es muy importante para conseguir la independencia.

Para probar si los residuales no están auto correlacionados, puede usarse la prueba Durbin-Watson, considerando el supuesto que los residuales corresponden a un proceso auto regresivo de primer orden.

La prueba para la presencia de autocorrelación en los residuales determinando si la correlación entre dos errores adyacentes es cero. La ecuación es:

$$d = \frac{\sum_{t=2}^N (e_t - e_{t-1})^2}{\sum_{t=1}^N e_t^2} \quad (4.44)$$

La autocorrelación está dada por la determinación del estadístico  $d$ , comparándolo con dos límites para llegar a una conclusión. Si  $d < d_L$  existe autocorrelación. Si  $d > d_U$  no existe autocorrelación. Si  $d_L \leq d \leq d_U$  la prueba no es concluyente.

Una vez cumplidos estos supuestos, se utilizan las extensiones de las pruebas de hipótesis univariadas para aplicarse a casos multivariados.

#### a) Prueba de la variabilidad del modelo

Primero se considera que ninguna de las  $x$  predice ninguna de las  $y$ . Esto puede ser expresado como:  $H_0: \mathbf{W}_1 = 0$ , donde  $\mathbf{W}_1$  incluye todos los renglones de  $\mathbf{W}$  excepto el bias:

$$\mathbf{W} = \begin{pmatrix} \mathbf{w}'_b \\ \mathbf{W}_1 \end{pmatrix} = \begin{pmatrix} w_{b1} & w_{b2} & \cdots & w_{bp} \\ w_{11} & w_{12} & \cdots & w_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ w_{q1} & w_{q2} & \cdots & w_{qp} \end{pmatrix}$$

No se incluye  $\mathbf{w}'_b = 0'$  en la hipótesis para no restringir las salidas a tener un bias igual a cero. La hipótesis alternativa es  $H_1: \mathbf{W}_1 \neq 0$ .

Es posible particionar la suma de cuadrados total y matriz de productos:

$$\mathbf{Y}'\mathbf{Y} = (\mathbf{Y}'\mathbf{Y} - \widehat{\mathbf{W}}'\mathbf{G}'\mathbf{Y}) + \widehat{\mathbf{W}}'\mathbf{G}'\mathbf{Y} \quad (4.45)$$

Así como en el caso univariado, se resta  $n\bar{y}\bar{y}'$  de ambos lados para evitar incluir  $\mathbf{w}'_b = \mathbf{0}'$ :

$$\mathbf{Y}'\mathbf{Y} - n\bar{y}\bar{y}' = (\mathbf{Y}'\mathbf{Y} - \widehat{\mathbf{W}}'\mathbf{G}'\mathbf{Y}) + (\widehat{\mathbf{W}}'\mathbf{G}'\mathbf{Y} - n\bar{y}\bar{y}') = \mathbf{E} + \mathbf{H} \quad (4.46)$$

Las notaciones  $\mathbf{H} = \widehat{\mathbf{W}}' \mathbf{G}' \mathbf{Y} - n\bar{y}\bar{y}'$  y  $\mathbf{E} = \mathbf{Y}' \mathbf{Y} - \widehat{\mathbf{W}}' \mathbf{G}' \mathbf{Y}$  pueden utilizarse para probar  $H_0: \mathbf{W}_1 = 0$  por medio de:

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{E} + \mathbf{H}|} = \frac{|\mathbf{Y}' \mathbf{Y} - \widehat{\mathbf{W}}' \mathbf{G}' \mathbf{Y}|}{|\mathbf{Y}' \mathbf{Y} - n\bar{y}\bar{y}'|} \quad (4.47)$$

El cual está distribuido como  $\Lambda_{p,q,n-q-1}$  cuando  $H_0: \mathbf{W}_1 = 0$  es verdad,  $p$  es el número de  $y$  y  $q$  el número de  $x$ . Se rechaza  $H_0$  si  $\Lambda \leq \Lambda_{p,q,n-q-1, \nu_H = q}$  y  $\nu_E = n - q - 1$  son los grados de libertad para  $\mathbf{H}$  y  $\mathbf{E}$  respectivamente.

#### b) Pruebas sobre subconjuntos de $x$

Se considera la hipótesis que las  $y$  no dependen de las  $h$  últimas de  $x$ ,  $x_{q-h+1}, x_{q-h+2}, \dots, x_q$ . Para esto se toma en cuenta que ninguna de las  $y_p$  es predicha por ninguna de esas  $x_h$ . Para expresar esta hipótesis, escribimos la forma particionada de  $\mathbf{W}$ .

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_r \\ \mathbf{W}_e \end{pmatrix}$$

Así como en el caso univariado, el subíndice  $r$  denota el subconjunto de  $w_{jk}$  conservados en el modelo reducido y  $d$  representa el subconjunto  $w_{jk}$  eliminado si no resulta ser significantes en la predicción de  $y$ . De esta forma  $\mathbf{W}_e$  tiene  $h$  renglones. La hipótesis puede ser expresada como:

$$H_0: \mathbf{W}_e = 0$$

Si  $\mathbf{G}_r$  contiene las columnas de  $\mathbf{W}$  correspondientes a  $\mathbf{W}_r$ , el modelo reducido es:

$$Y = G_r W_r + \Xi \quad (4.48)$$

Para comparar el ajuste del modelo completo y el modelo reducido, se utiliza la diferencia entre la matriz de productos y suma de cuadrados del modelo para ambos modelos. Esta diferencia se convierte en la nueva matriz  $H$ .

$$H = \widehat{W}' G' Y - W_r' G_r' Y \quad (4.49)$$

Para realizar esta prueba, se utiliza la matriz  $E$  basado en el modelo completo:

$$E + H = (Y'Y - \widehat{W}' G' Y) + (\widehat{W}' G' Y - W_r' G_r' Y) = Y'Y - W_r' G_r' Y \quad (4.50)$$

Y el estadístico lambda de Wilks' está dado por:

$$\Lambda(x_{q-h+1}, \dots, x_q | x_1, \dots, x_{q-h}) = \frac{|E|}{|E + H|} = \frac{|Y'Y - \widehat{W}' G' Y|}{|Y'Y - W_r' G_r' Y|} \quad (4.51)$$

El cual está distribuido como  $\Lambda_{p,h,n-q-1}$  cuando  $H_0: W_e = 0$  es verdad, con  $v_H = q$  y  $v_E = n - q - 1$  grados de libertad.

Las ecuaciones anteriores se resumen en la Tabla 4.1.

**Tabla 4.1 Construcción de la tabla MANOVA**

Fuente de variación	Grados de libertad	Suma de cuadrados	$\Lambda$
Modelo	$q$	$ H $	$\frac{ E }{ E + H }$
Residual	$n - q - 1$	$ E $	
Total	$n - 1$	$ E + H $	

#### 4.12 Elipsoides de confianza e intervalos

Además de las estimaciones puntuales del bias y los pesos, es posible obtener los intervalos de confianza de las estimaciones de estos parámetros. El ancho de estos intervalos de confianza es una medida de la calidad global del modelo RBF.

Si los términos de error, en el modelo se distribuyen normalmente y de forma independiente, se distribuyen como una F. Esto lleva a la siguiente definición de  $100(1 - \alpha)\%$  de intervalo de confianza en los pesos.

$$w_{jq} \pm \sqrt{\frac{(n-p)(p-1)}{n-p-q} F_{q,n-p-q}(\alpha) \hat{\sigma}_{ii}} \quad (4.52)$$

Una aplicación importante del modelo es la predicción de nuevas observaciones de  $Y$ , correspondiente a un nivel específico de la variable  $x$ . Si es de interés predecir una respuesta individual en un espacio  $q$ -dimensional, el punto predictor de  $Y_0$  es  $g'_0 \hat{w}$ , y el error de predicción  $Y_0 - g'_0 \hat{w} = (w - g'_0 \hat{w}) + e_0$  es distribuido como  $N_q(0, (1 + g'_0(G'G)^{-1}g_0)\Sigma)$ , entonces, el elipsoide de predicción al  $100(1 - \alpha)\%$  consiste en todos los valores de  $Y_0$  tal que:

$$\begin{aligned} & (Y_0 - g'_0 \hat{w})' \left( \frac{n}{n-p-1} \hat{\Sigma} \right)^{-1} (Y_0 - g'_0 \hat{w}) \\ & \leq (1 + g'_0(G'G)^{-1}g_0) \left[ \left( \frac{q(n-p-1)}{n-p-q} \right) F_{q,n-p-q}(\alpha) \right] \end{aligned} \quad (4.53)$$

Simultáneamente, los intervalos de predicción para las  $q$  variables de respuesta son:

$$(4.54)$$

$$g'_{0} \widehat{w}_{(i)} = \sqrt{\left(\frac{q(n-p-1)}{n-p-q}\right) F_{q,n-p-q}(\alpha)} \sqrt{(1 + g'_{0}(G'G)^{-1}g_{0}) \left(\frac{n}{n-p-1} \widehat{\sigma}_{ii}\right)}$$

Donde  $\widehat{w}_{(i)}$  es la  $i$ -ésima columna de la matriz de pesos  $\widehat{w}$ , y  $\widehat{\sigma}_{ii}$  es el  $i$ -ésimo elemento de la diagonal de  $\widehat{\Sigma}$ .

#### 4.13 Medidas de asociación y validación de la predicción

La ecuación (4.27) usada como función de evaluación del Algoritmo Genético. En el caso univariado es una medida de la reducción proporcional de la variación total en la variable dependiente y usando un conjunto fijo de variables independientes  $x_1, x_2, \dots, x_p$ . Una extensión de  $R^2$  en el caso multivariado es la relación de correlación de Fisher ( $\eta^2$ ) propuesto por Wilk. En el caso multivariado, eta cuadrado es llamado el cuadrado del vector de coeficientes de correlación definido por:

$$\eta^2 = 1 - \Lambda \quad (4.55)$$

Donde  $\Lambda$  está definido por (4.45). La métrica  $\eta^2$  es una medida sesgada, para utilizar una medida insesgada se propone:

$$\eta_a^2 = 1 - \frac{n\Lambda}{(n-q+\Lambda)} \quad (4.56)$$

Otra medida de asociación es  $\eta_{\theta}^2$ , el cual está basado en el cuadrado de la mayor correlación canónica. Esta dada por (4.56).

$$\eta_{\theta}^2 = \frac{\lambda_1}{1+\lambda_1} = \theta_1 \leq \eta^2 \quad (4.57)$$

Una estimación global de la predicción para variables fijas independientes es:

$$\eta_c^2 = 1 - \left( \frac{n + q + 1}{n} \right) (1 - \eta_a^2) \quad (4.58)$$

Como una extensión del modelo multivariado, es posible considerar el estimador:

$$\eta_{cb}^2 = \frac{(n - q - 3)(\eta_a^2)^2 + \eta_a^2}{(n - 2q - 2)\eta_a^2 + q} \quad (4.59)$$

Para  $q$  parámetros o  $p$  variables independientes en el modelo multivariado, (4.28) es una matriz cuadrada ( $q \times q$ ) y es una extensión directa de (4.27), y puede utilizarse también como medida de asociación.

Para evaluar el ajuste de  $Y$  a  $\hat{Y}$ , es posible estimar cada fila  $y_i'$  de  $Y$  usando  $\hat{y}'_{i(i)} = x_i' \widehat{W}_{q(i)}$  donde el peso  $\widehat{W}_{q(i)}$  es estimado eliminando la  $i$ -ésima fila de  $y$  y  $X$  para varios valores de  $q$ . La cantidad  $y_i - \hat{y}_{i(i)}$  es llamado el residual eliminado y sumando los productos de todas las observaciones  $i = 1, 2, \dots, n$  obteniendo el criterio de la suma de cuadrados de la predicción multivariada:

$$MPRESS_q = \sum_{i=1}^n (y_i - \hat{y}_{i(i)})' (y_i - \hat{y}_{i(i)}) = \sum_{i=1}^n \frac{\hat{e}_i' \hat{e}_i}{(1 - p_{ii})^2} \quad (4.60)$$

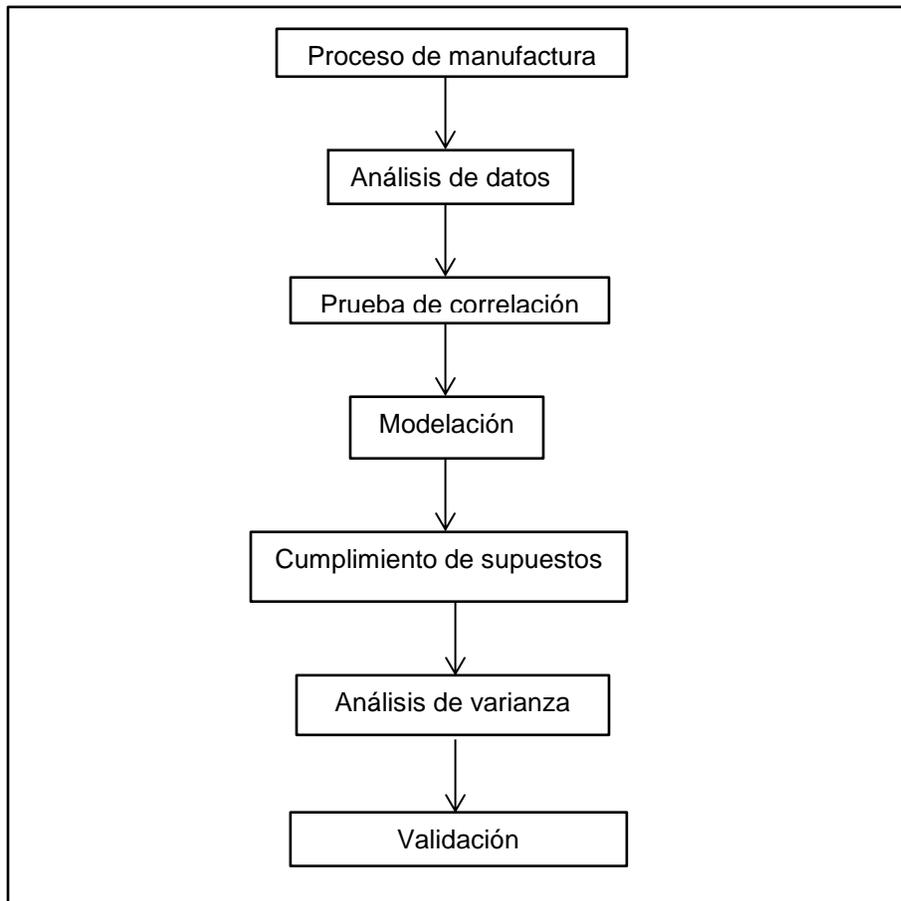
Los modelos multivariados con un  $MPRESS_q$  pequeño, son considerados en la selección.

El coeficiente  $R_{MPRESS_q}^2$  consiste en el estadístico  $MPRESS_q$ , definido por (4.59) y la respuesta del experimento representado por  $Y_i$ , está definido por (4.60).

$$R_{pred_q}^2 = 1 - \frac{MPRESS_q}{tr(E + H)} \quad (4.61)$$

## Capítulo 5

### Metodología



**Figura 5.1 Metodología general**

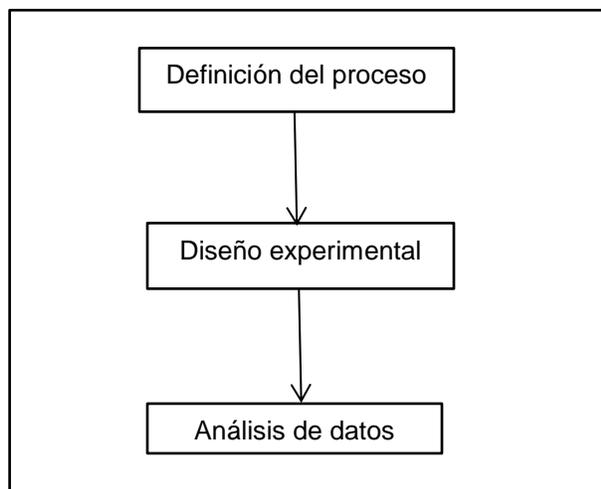
En el presente capítulo se abordará la metodología propuesta para tratar el problema. La Figura 5.1 detalla la metodología propuesta de manera general.

- **Definición del proceso de manufactura**

Se tiene la necesidad de modelar un proceso de manufactura que trabaja con especificaciones de entrada para generar características de calidad en el producto, y explicar su comportamiento mediante la red de Función de Base Radial.

- **Diseño de experimentos**

En esta etapa, se determinan qué factores intervienen en el comportamiento de una variable de respuesta, por medio de una fase de experimentación o la referencia de datos históricos. Las primeras dos etapas se describen en la Figura 5.2.



**Figura 5.2 Análisis de datos**

- **Correlación de respuestas**

Si se modelan varias respuestas, se verifica si están **correlacionadas** significativamente, mediante el coeficiente de correlación de Pearson descrito por la ecuación (4.38), recordando:

$$\rho_{ij} = \frac{\text{cov}(Y_i, Y_j)}{\sqrt{\text{var}(Y_i)\text{var}(Y_j)}}$$

Para determinar si las variables están significativamente correlacionadas, se utilizan las tablas de valores críticos para el coeficiente de correlación Pearson. Si el valor de correlación es mayor al valor crítico, se dice que las variables están significativamente correlacionadas. Otra opción es por medio de una prueba de hipótesis al determinar el *P – value* del coeficiente Pearson. Las hipótesis se definen de la siguiente manera:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

El rechazo de la hipótesis nula indica correlación significativa.

Una vez determinada la correlación significativa entre respuestas, se procede a modelar y analizar el proceso de forma multivariada. Mientras que, si las variables de respuesta no están correlacionadas o solo se tiene una respuesta, el modelo y análisis del proceso será univariado.

- **Modelación**

El siguiente paso en la metodología propuesta es la modelación, en este paso se utiliza una red de base radial dependiendo de las respuestas del proceso. Las salidas se representan por la ecuación (4.19), que es la siguiente:

$$y(x) = \sum_{i=1}^n w_i \mathbf{G}(\|x - c_i\|) + b + \varepsilon$$

Para la determinación de los centros  $c_i$  se utiliza un algoritmo genético, cuya función objetivo será  $R_q^2$  definido por (4.27), para el caso univariado y multivariado.

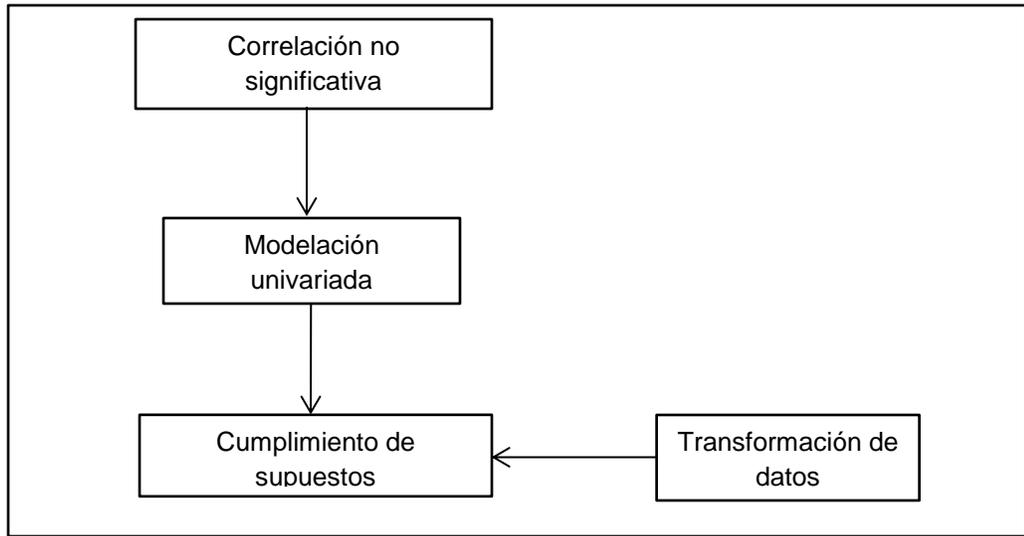
La estimación de los pesos  $w_i$  también será definida por las respuestas del modelo, usando el estimador de mínimos cuadrados (ecuación (4.24)).

- **Cumplimiento de supuestos**

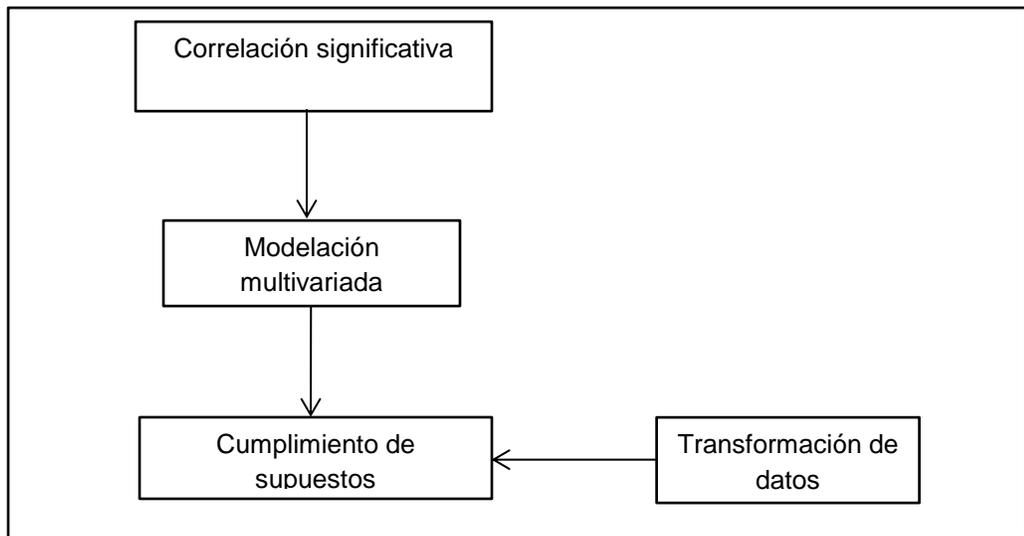
En el siguiente paso se verifica si el modelo cumple los supuestos para el análisis de varianza correspondiente, en este caso MANOVA. A continuación, se presentan los supuestos que se deben cumplir y su método de verificación:

- a) Las respuestas deben estar normalmente distribuidos de forma conjunta, prueba de Mardia multivariada ecuaciones (4.33) y (4.34).
- b) Las respuestas deben estar correlacionadas linealmente, prueba de Pearson ecuación (4.38).
- c) Las respuestas deben tener varianzas iguales, prueba de White ecuación (4.42).
- d) Las respuestas deben estar distribuidas independientemente, prueba Durbin-Watson ecuación (4.44).

Si no se cumplen estos supuestos, se verifica si es posible realizar alguna transformación en las variables para cumplirlos. Las Figuras 5.3 y 5.4 describen los pasos de la determinación de la correlación y la modelación en caso univariado y multivariado.



**Figura 5.3 Correlación no significativa**



**Figura 5.4 Correlación significativa**

- **Análisis de varianza**

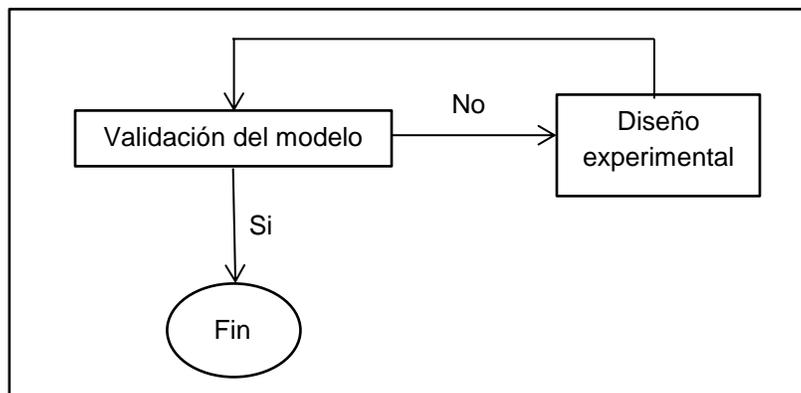
Una vez verificados los supuestos, el siguiente paso es realizar el análisis para contestar las siguientes preguntas:

- ¿Qué tan bien estará explicada la variabilidad relacionada con el proceso?
- ¿Cómo sabemos que las variables incluidas en la red tienen influencia significativa en las respuestas?

Para responder estas preguntas se utiliza el análisis de varianza correspondiente (ANOVA o MANOVA).

- **Validación**

La última etapa de la metodología es realizar una validación, tanto en campo como utilizando indicadores estadísticos como  $R_q^2$  o  $\eta^2$ . Si en esta etapa no se obtienen los resultados deseados, se repetirá el diseño experimental y se volverá a modelar hasta obtener una representación adecuada del proceso que se está analizando. La Figura 5.5 muestra el diagrama de la validación.



**Figura 5.5 Diagrama de validación**

En el desarrollo de la presente investigación, se tiene como principal aplicación una red neuronal de Función de Base Radial usada en predicción para varias respuestas en un proceso de fundición en molde permanente y por gravedad, analizada mediante técnicas estadísticas multivariadas.

## Capítulo 6

### Aplicación

La metodología propuesta en el capítulo anterior, se propone desde un punto de vista general. Una vez definido el proceso a analizar, los pasos siguientes son los expuestos en el capítulo anterior. La Figura 6.1 retoma los pasos de la metodología propuesta y en el presente capítulo se detalla la metodología paso a paso y los resultados obtenidos en el caso de estudio.

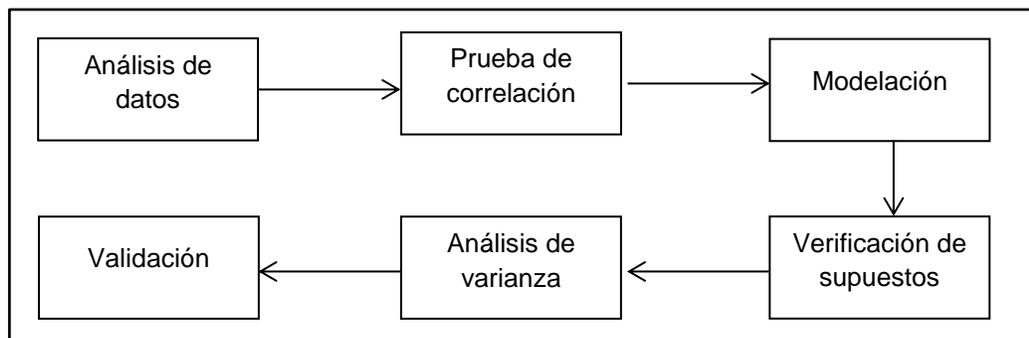


Figura 6.1 Metodología propuesta

#### 6.1 Descripción del proceso de manufactura

La metodología se aplicó en la planta de fundición Aluminio y Bronces de Saltillo S.A. de C.V. (ALBRONSA). En el proceso se vacían piezas de aleación ZAMAK 3, denominada así por la norma ASM de aleaciones por su composición química. Esta aleación se utiliza en la fabricación de la masa para cabecera de asiento para automóvil.

El proceso de vaciado es en molde permanente y consta de cuatro pasos principales: 1) precalentar y recubrir el molde; 2) insertar corazones y cerrar el molde; 3) vaciar el metal fundido al molde y 4) abrir el molde y sacar la pieza. (Dieter, 1988).

Para este análisis se utilizaron como variables de entrada:

- $x_1$ : temperatura del metal
- $x_2$ : temperatura del molde
- $x_3$ : basculamiento

Estas variables de entrada afectan a las características de calidad deseadas, como el peso de la pieza, SPC 1, SPC 2, SPC 3 y SPC 4 (Figura 6.2). Por las condiciones del proceso, se sabe que el peso de la pieza tiene que ver con defectos como mal llenado o porosidad, SPC 2 con rechupes y SPC 3 con desprendimiento de colada o mal llenado. Por otro lado, SPC 1 y SPC 4 tienen que ver con el acomodo de la pieza en el molde. Por lo tanto, el análisis tiene por objetivo disminuir la variabilidad del proceso en las respuestas peso, SPC 2 y SPC 3. La Figura 6.2 muestra la pieza fabricada y las características de calidad medidas.

Los datos del proceso que se tomaron y analizaron se presentan en la siguiente sección.

## 6.2 Análisis de datos

Los datos experimentales se obtuvieron mediante un diseño central compuesto de un diseño factorial completo  $2^3$ , seis puntos axiales elegidos con un  $\alpha = 1.682$  y seis puntos centrales para un total de 20 observaciones mostradas en la Tabla 6.1.

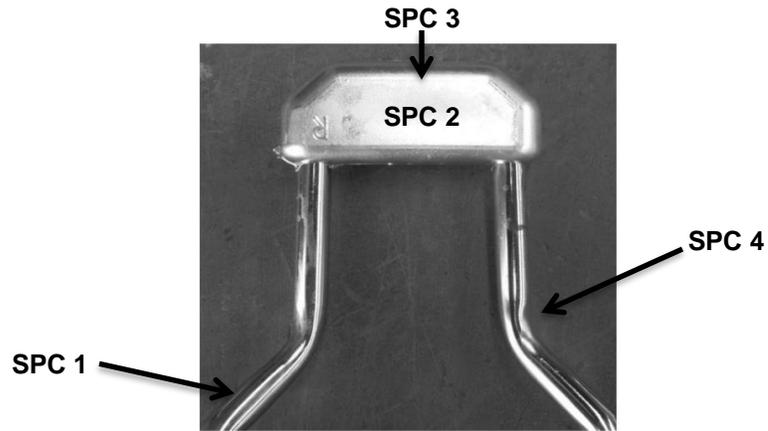


Figura 6.2 Pieza fabricada

Tabla 6.1 Datos del proceso

$x_1$ : Temperatura del metal (°C)	$x_2$ : Temperatura del molde (°C)	$x_3$ : Basculamiento (seg)
390	210	18
420	210	18
390	250	18
420	250	18
390	210	30
420	210	30
390	250	30
420	250	30
379.77	230	24
430.22	230	24
405	196.36	24
405	263.63	24
405	230	13.90
405	230	34.09
405	230	24
405	230	24
405	230	24
405	230	24
405	230	24
405	230	24
405	230	24
405	230	24

El objetivo de la modelación es disminuir la variabilidad del proceso y encontrar la variable que causa más defectos y minimizarlos. Tales defectos del proceso son principalmente poro, poro de gas, rechupe, desprendimiento de colada y marcas, estos están relacionados con las características de calidad peso, SPC 2 y SPC 3, los cuales fueron tomados como respuestas del proceso.

Sin embargo, el molde de vaciado del proceso cuenta con seis cavidades M, N, O, P, Q, R de las cuales cada cavidad es una pieza. Aquí surge la necesidad de conocer si existe alguna diferencia en las características de calidad. Para esto, se propone una prueba de Bartlett de homogeneidad de varianzas, utilizando los seis puntos centrales del DOE. Se rechaza  $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$  cuando al menos una varianza es diferente.

Primero se analizó la respuesta  $y_1$ : **peso**, con:  $x_1 = 402.5$ ,  $x_2 = 230$  y  $x_3 = 24$ , los datos se muestran en la Tabla 6.2.

**Tabla 6.2 Puntos centrales de  $y_1$**

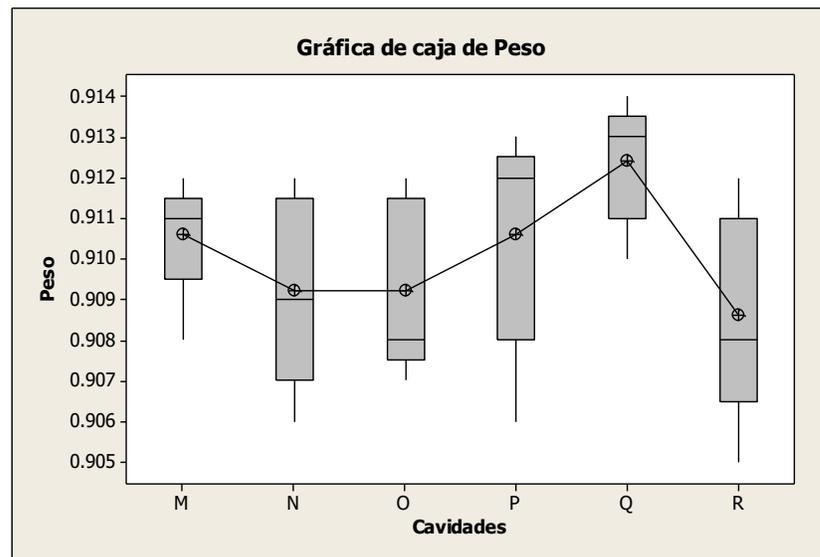
M	N	O	P	Q	R
0.911	0.912	0.907	0.910	0.912	0.908
0.908	0.908	0.912	0.912	0.913	0.908
0.911	0.906	0.908	0.913	0.913	0.912
0.911	0.911	0.911	0.906	0.910	0.905
0.912	0.909	0.908	0.912	0.914	0.910
0.875	0.906	0.908	0.879	0.911	0.912

Se realizó la prueba de Bartlett estimando un estadístico  $T_B = 2.37398$ , y un valor de probabilidad de  $P = 0.79534$ . El valor crítico de tablas  $\chi^2$  con 4 grados de libertad y un nivel de significancia de 0.05 es  $\chi_{5,0.05}^2 = 9.49$ . Se rechaza  $H_0$  bajo dos criterios, si  $T_B > \chi_{4,0.05}^2$  y  $0.05 > P$ . Con esta información es posible afirmar que las variancias de las muestras son homogéneas. Para verificar qué cavidad tiene mayor varianza se utiliza la Tabla 6.3 y la Figura 6.3

**Tabla 6.3 Medias de cada cavidad de  $y_1$**

Nivel	N	Media	Desv.Est.	ICs de 95% individuales para la media basados en Desv.Est. agrupada
M	6	0.90433	0.01314	(-----*-----)
N	6	0.90867	0.00250	(-----*-----)
O	6	0.90900	0.00200	(-----*-----)
P	6	0.90467	0.01460	(-----*-----)
Q	6	0.91217	0.00147	(-----*-----)
R	6	0.90917	0.00271	(-----*-----)

-----+-----+-----+-----+-----  
0.9000    0.9060    0.9120    0.9180



**Figura 6.3 Gráfica de caja para  $y_1$**

Con la información anterior se observa que la cavidad con mayor variación es la cavidad P, ya que tiene una mayor desviación estándar y un tamaño de caja mayor.

Este análisis se repitió para las dos respuestas restantes y se muestran a continuación. En el caso de  $y_2$ : *SPC 2* el valor crítico de tablas es  $\chi_{5,0.05}^2 = 11.07$ . Se utilizaron los datos de la Tabla 6.4.

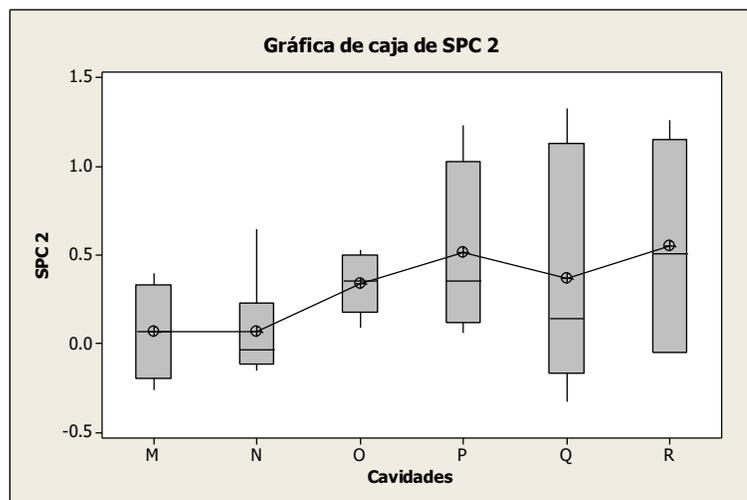
**Tabla 6.4 Puntos centrales de  $y_2$**

M	N	O	P	Q	R
0.23	0.65	0.47	0.96	0.20	0.77
0.40	0	0.21	0.26	-0.33	-0.05
0.31	0.09	0.53	0.14	-0.11	0.24
-0.17	-0.15	0.09	0.44	1.06	1.26
-0.10	-0.1	0.24	0.06	0.08	-0.05
-0.26	-0.07	0.49	1.23	1.33	1.12

Con estos datos se obtuvo un estadístico de Bartlett  $T_B = 10.0595$  y un valor  $P = 0.0736$ , rechazando así  $H_0$  por ambos criterios. Se utiliza información de la Tabla 6.5 y la Figura 6.4 para seleccionar la cavidad de mayor variación.

**Tabla 6.5 Medias de cada cavidad de  $y_2$**

Nivel	N	Media	Desv.Est.	ICs de 95% individuales para la media basados en Desv.Est. agrupada
M	6	0.0683	0.2784	(-----*-----)
N	6	0.0700	0.2962	(-----*-----)
O	6	0.3383	0.1816	(-----*-----)
P	6	0.5150	0.4749	(-----*-----)
Q	6	0.3717	0.6680	(-----*-----)
R	6	0.5483	0.5820	(-----*-----)



**Figura 6.4 Gráfica de caja para  $y_2$**

Por último, para seleccionar la cavidad de la respuesta  $y_3$ : **SPC 3** se utilizan los datos de la Tabla 6.6.

**Tabla 6.6 Puntos centrales de  $y_3$**

M	N	O	P	Q	R
0.86	1.52	0.45	0.36	1.11	0.80
0.70	0.70	0.37	0.35	0.40	0.69
0.95	1.05	0.29	0.43	0.70	0.63
0.69	1	0.62	0.47	1.10	0.89
0.72	0.91	0.42	0.10	0.86	0.42
0.58	1.15	0.43	0.51	1.15	0.92

En este caso  $T_B = 7.35866$  y  $P = 0.1953$ , rechazando así la hipótesis de igualdad de varianzas también en este caso. Para conocer la cavidad con mayor variación se utiliza la información de la Tabla 6.7.

Se puede observar gráficamente una desviación estándar alta en la cavidad Q. La variación se muestra en el tamaño de caja en la Figura 6.5.

**Tabla 6.7 Medias de cada cavidad de  $y_3$**

Nivel	N	Media	Desv.Est.	ICs de 95% individuales para la media basados en Desv.Est. agrupada
M	6	0.7500	0.1327	(-----*-----)
N	6	0.8867	0.2957	(-----*-----)
O	6	0.4300	0.1094	(-----*-----)
P	6	0.3700	0.1460	(-----*-----)
Q	6	1.0550	0.2740	(-----*-----)
R	6	0.7250	0.1866	(-----*-----)

-----+-----+-----+-----+-----  
0.30      0.60      0.90      1.20

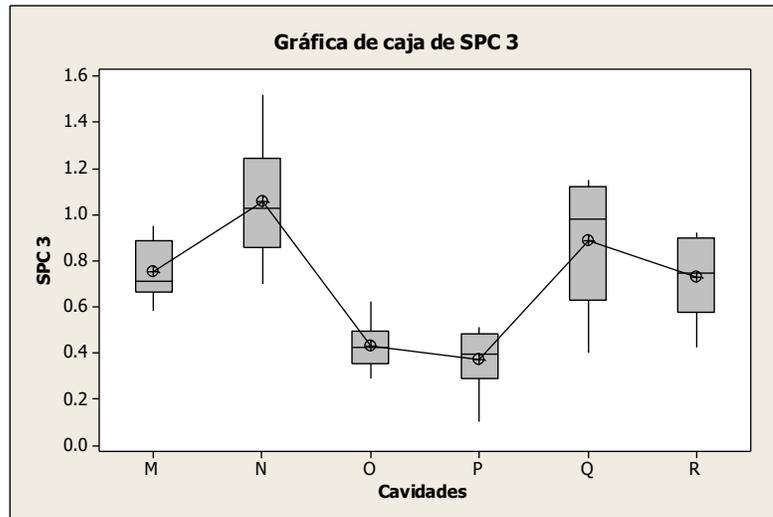


Figura 6.5 Gráfica de caja para  $y_3$

Por lo tanto, se elige la cavidad Q para modelar ya que tiene mayor variación en las respuestas. Los datos se muestran en la Tabla 6.8.

El siguiente paso en la metodología es determinar si existe correlación entre las respuestas seleccionadas, para determinar si se modela univariado o multivariado. La prueba de correlación se muestra en la sección siguiente.

### 6.3 Prueba de correlación

Una vez obteniendo los datos del proceso, el siguiente paso en la metodología propuesta es verificar si existe correlación significativa entre las respuestas y determinar si se modela de forma univariada o multivariada. En este paso se utiliza el coeficiente de correlación de Pearson mostrado en la ecuación (4.38) y su valor de probabilidad. Los resultados se muestran en la Tabla 6.9.

**Tabla 6.8 Variables de respuesta**

$y_1$ : <b>Peso (kg)</b>	$y_2$ : <b>SPC 2 (mm)</b>	$y_3$ : <b>SPC 3 (mm)</b>
0.913	0.93	0.98
0.917	1.00	1.00
0.916	0.10	0.79
0.918	0.80	0.70
0.919	0.80	0.78
0.910	0.20	1.11
0.912	-0.33	0.40
0.913	-0.11	0.70
0.906	1.06	1.10
0.912	0.08	0.86
0.879	1.33	1.15
0.910	0.02	0.80
0.914	0.28	0.84
0.934	-0.40	0.53
0.910	-0.20	0.73
0.906	0.05	0.80
0.907	0.12	0.98
0.912	-0.06	0.75
0.920	-0.84	0.77
0.920	0.58	0.25

**Tabla 6.9 Coeficientes de correlación de Pearson**

	$y_1$ : <i>peso</i>	$y_2$ : <i>SPC 2</i>
$y_2$ : <i>SPC 2</i>	$\rho = -0.433$ $Pvalue = 0.057$ $VC = 0.378$	– –
$y_3$ : <i>SPC 3</i>	$\rho = -0.571$ $Pvalue = 0.009$ $VC = 0.378$	$\rho = 0.459$ $Pvalue = 0.042$ $VC = 0.378$

En la Tabla (6.9) se observa que el valor de probabilidad  $P$  es mayor al nivel de significancia  $\alpha = 0.05$  solo en la relación  $y_1$ : *peso* y  $y_2$ : *SPC 2* pero es muy cercano al  $\alpha$ . En este caso se utiliza el criterio del valor crítico para el coeficiente de correlación, el cual es de 0.378, y debido a que el coeficiente supera este valor crítico estas variables se consideran correlacionadas. Los demás coeficientes también superan el valor crítico y cumplen con la condición de ser menores en su valor  $P$  a 0.05.

Por otra parte, el coeficiente de correlación no paramétrico de Kendall es menos sensible a la correlación lineal, se recomienda su uso debido al comportamiento no lineal común de los procesos de manufactura, es probable que las variables de respuesta tengan una correlación no lineal. La Tabla 6.10 muestra los coeficientes de correlación de Kendall mostrado en la ecuación (4.39) para las variables de respuesta.

Esta prueba no indica correlación entre  $y_1$ : *peso* y  $y_2$ : *SPC 2*. Sin embargo, ambas variables muestran relación significativa con  $y_3$ : *SPC 3* ya que su valor de probabilidad  $P$  es menor a 0.05, por lo tanto, se recomienda trabajar con las tres variables de manera conjunta.

En base a los resultados con las dos pruebas, se determina que las respuestas tienen correlación significativa y se propone modelar el proceso mediante técnicas multivariadas, como se propone a continuación.

**Tabla 6.10 Coeficientes de correlación de Kendall**

	$y_1$ : <i>peso</i>	$y_2$ : <i>SPC 2</i>
$y_2$ : <i>SPC 2</i>	$\tau = -0.1189$ $Pvalue = 0.4931$	– –
$y_3$ : <i>SPC 3</i>	$\tau = -0.4348$ $Pvalue = 0.0098$	$\tau = 0.5000$ $Pvalue = 0.0025$

## 6.4 Modelación

Los centros de la red se calcularon mediante un algoritmo genético y utilizando como función de evaluación la extensión multivariada del coeficiente de determinación. Maximizando la función descrita por la ecuación (4.27), recordando:

$$R_q^2 = (Y'Y - n\bar{y}\bar{y}')^{-1}(\widehat{B}'X'Y - n\bar{y}\bar{y}')$$

Dicha solución se buscó mediante un algoritmo de representación binaria de posibles soluciones, donde cada carácter representa un bit. Se definió una población inicial de 100 individuos de 16 bits, una selección de torneo de tamaño 2 y una mutación simple, una probabilidad de cruce de 0.8 y de mutación de 0.2, generando ciclos durante 100 generaciones. Los centros generados se presentan en la Tabla 6.11.

**Tabla 6.11 Centros generados por algoritmo genético**

419.34	224.28	28.8519
403.98	215.19	22.2385
402.89	238.76	28.5127

La distancia entre las variables de entrada del proceso y los centros, se calculó mediante la distancia de Mahalanobis calculada con la ecuación (4.28) donde  $x$  son las variables del proceso,  $c$  los centros calculados mediante algoritmo genético y  $\Sigma^{-1}$  es la inversa de la matriz de covarianza.

Así, la estructura de la red para esta aplicación incluye tres nodos en la capa de entrada, tres en la capa oculta y tres en la capa de salida. La función utilizada fue la ecuación (4.12) por obtener el mejor resultado en la función de evaluación del algoritmo genético.

Los pesos  $w$  son determinados mediante (4.24), el estimador de mínimos cuadrados multivariados:

$$\widehat{W} = (G'G)^{-1}G'Y$$

La matriz de pesos se presenta en la Tabla 6.12.

**Tabla 6.12 Pesos generados por mínimos cuadrados**

Bias	0.9122	0.5693	0.8453
$w_1$	-0.0007	-1.0796	0.5106
$w_2$	-0.0558	2.6264	1.1069
$w_3$	0.0656	-4.6627	-1.8751

Por medio de la ecuación (6.1) se estiman las salidas de la red. El siguiente paso es verificar si se cumplen los supuestos para el análisis multivariado de varianza. Este paso de la metodología se realiza en la sección siguiente.

La ecuación (4.19) muestra la relación de entradas y salidas de la red, el modelo estaría representado por:

$$y(x) = \sum_{i=1}^n w_i G(\|x - c_i\|) + b + \varepsilon$$

Donde  $w_i$  son los pesos de la red determinados por la Tabla 6.12 del renglón dos al cuatro,  $x$  por la Tabla 6.2,  $c_i$  son los centros por la Tabla 6.11, y  $b$  el bias por la Tabla 6.12 en el renglón uno.

## 6.5 Cumplimiento de supuestos

Una vez construida la red, el siguiente paso es comprobar los supuestos para el análisis de varianza, en este caso multivariado, los cuales son los siguientes:

### 1. Normalidad multivariada

Se utilizó la prueba Henze-Zirkler para determinar la normalidad entre las respuestas. Mediante la ecuación (4.29) se calculó el estadístico Henze-Zirkler el cual se presenta a continuación:

$$T_{\beta} = 1.1420$$

Y su valor  $P$  asociado es:

$$P - value = 0.0015$$

Con un nivel de significancia del 0.05, la hipótesis que indica que los datos son normales se rechaza ya que el valor de probabilidad es menor al nivel de significancia, por lo tanto, se dice que los datos no son normales.

Sin embargo, la literatura indica que se deben realizar varias pruebas para determinar la normalidad de los datos (Mardia 1980, Mecklin y Mundfrom 2003). Por lo tanto, se utilizó la prueba de simetría y curtosis de Mardia para determinar si las respuestas siguen una distribución normal multivariada. Aplicando las ecuaciones (4.33) y (4.34), los coeficientes de simetría y curtosis determinados y sus valores  $P$  son:

$$\hat{\beta}_{1,p} = 19.06$$

$$P - value = 0.0394$$

$$\hat{\beta}_{2,p} = 1.6015$$
$$P - value = 0.1093$$

Estos coeficientes indican que los datos tienen simetría significativa, mientras que el coeficiente de curtosis indica datos dispersos al no ser significativa. Dicha dispersión puede ser la causa de que otras pruebas de bondad de ajuste determinen no normalidad en estos datos. Sin embargo, la distribución normal es la única distribución simétrica. Por lo tanto, aunque la prueba indique dispersión, las respuestas se consideran normales.

Otra prueba es el estadístico de prueba Shapiro-Wilk, el cual compara el valor obtenido con un valor crítico de tablas. Se rechaza la hipótesis de normalidad si el valor obtenido es menor al valor de tablas.  $W_M$  se calcula mediante la ecuación (4.37), se muestra dicho valor y el valor crítico:

$$W_M = 0.9068$$
$$VC = 0.9050$$

Es posible observar que el valor obtenido es mayor al valor de tablas, por lo tanto, los datos se consideran normales.

## 2. Linealidad

Para el segundo supuesto se utilizó el coeficiente de correlación Pearson, expuesto en la Tabla 6.3 y el coeficiente de correlación de Kendall expuesto en la Tabla 6.4.

### 3. Homogeneidad de varianzas intra-vectores

El supuesto de homogeneidad de varianzas indica que las varianzas entre los residuales de los grupos de variables deben ser homogéneas. Dicho supuesto se verifica mediante la prueba White de homocedasticidad. El estadístico de Homocedasticidad de la prueba de White está definida por la ecuación (4.42), recordando:

$$V = \sigma^2[X'X]^{-1}$$

Los valores de esta prueba para cada grupo de variables son:

$$V_1 = 8.56$$

$$P = 0.0309$$

$$V_2 = 5.7502$$

$$P = 0.0029$$

$$V_3 = 2.1660$$

$$P = 0.0000021$$

Todos menores que el valor ji cuadrado de tablas para  $n - p - 1$  grados de libertad de 26.3, y un valor  $P$  menor al nivel de significancia 0.05. Esto indica el cumplimiento del supuesto de varianzas homogéneas.

### 4. Independencia intra-vectores

Mediante la prueba de Durbin-Watson se determina si los renglones de cada vector de respuestas no están auto-correlacionados. El estadístico que demuestra esto es la ecuación (4.32):

$$d = \frac{\sum_{t=2}^N (e_t - e_{t-1})^2}{\sum_{t=1}^N e_t^2}$$

Los valores para cada variable son:

$$d_1 = 1.6763$$

$$d_2 = 2.6131$$

$$d_3 = 1.5087$$

Mientras que los límites  $d_l$  y  $d_u$  con un nivel de significancia de 0.05 respectivamente son: 0.899 y 1.018. Así, es posible concluir que cada variable cumple con el supuesto de independencia serial al cumplirse que  $d > d_U$  en los tres casos.

Al cumplirse los supuestos necesarios para el análisis de varianza multivariado, se procede a la construcción de la tabla MANOVA, la cual se muestra en la sección siguiente.

## 6.6 Análisis de varianza multivariado

El primer análisis MANOVA, mostrado en la Tabla 6.13, indica la significancia estadística de la relación entre las variables dependientes y con una o más variables independientes  $x$ , así como la contribución de cada factor a la variación de las respuestas.

Como puede observarse en dicha tabla anterior, el valor Lambda del modelo multivariado es  $\Lambda = 0.1535$ . Dicho valor es comparado con los valores críticos de tablas para el Lambda de Wilk, en este modelo el valor crítico de tablas es  $\Lambda_{0.05,3,3,16} = 0.3340$ , para un alfa de 0.05, una dimensión de 3 por las respuestas que se tienen, y los grados de libertad del modelo y del residual de 3 y 16 respectivamente. Debido a que  $\Lambda < \Lambda_{0.05,3,3,16}$  se concluye que el modelo

explica bien la influencia de las variables independientes sobre las dependientes.

**Tabla 6.13 MANOVA para significancia del modelo**

<b>Fuente</b>	<b>Grados de libertad</b>	<b>Suma de cuadrados</b>	<b>Cociente</b>	<b>VC</b>
Modelo	3	0.000132	0.1535	$\Lambda_{0.05,3,3,16}$ = 0.3340
Residual	16	0.000895		
Total	19	0.0058		

Para conocer qué tan bien explica el modelo la variación del proceso, se estiman las métricas de evaluación multivariadas presentadas en las ecuaciones (4.55) -(4.59). Dichas métricas se presentan en la Tabla 6.14.

**Tabla 6.14 Medidas de asociación y predicción multivariadas**

$\eta^2$	84.85%
$\eta_{\alpha}^2$	82.10%
$\eta_c^2$	78.52%
$\eta_{cb}^2$	79.81%
$R_q^2$	74.59%

Debido a que las métricas mostradas en la Tabla 6.14 son mayores al 70% y no hay entre ellas una diferencia mayor a 10%, se determina que el modelo explica la variabilidad de manera confiable.

Por último, se determina qué variable es más importante para predecir las salidas deseadas. Esto mediante las pruebas de subconjuntos de  $x$  descritos en el capítulo 4 y resumidos en la Tabla 6.15.

Tabla 6.15 MANOVA para subconjuntos de  $x$

Fuente	Grados de Libertad	Suma de cuadrados	Cociente	VC
Temperatura de metal	2	8.8981	1.4640	$\Lambda_{0.05,2,2,17}$ = 0.5620
<b><u>Temperatura de molde</u></b>	<b><u>2</u></b>	<b><u>0.5620</u></b>	<b><u>0.4072</u></b>	
Basculamiento	2	11.5532	0.9041	
Residual	17	1.1730		
Total	19	79.0381		

En dicha tabla puede observarse el nivel de influencia de cada variable independiente sobre las dependientes, y concluir que la variable más importante en el proceso es la temperatura del molde, ya que es la única que está por debajo del valor crítico de 0.5620.

Con el análisis anterior se puede decir que el modelo propuesto de red neuronal de Función de Base Radial es adecuado y cumple con los supuestos del análisis multivariado de varianza. La predicción de los datos con la red se muestra en la Tabla 6.16.

## 6.7 Validación

Como parte de la validación de la red, se dividieron los datos en dos partes, unos para datos de estimación y otros de predicción. Esto mediante la métrica multivariada  $MPRESS_q$  calculando el siguiente vector:

$$MPRESS_q = \begin{bmatrix} 0.0020 \\ 3.3152 \\ 0.5669 \end{bmatrix}$$

Con estos datos es posible calcular un coeficiente  $R^2_{pred_q} = 0.8171$  por medio de la ecuación (4.61). Dado que el resultado es mayor a 0.8, se dice que el modelo es bueno para predecir.

**Tabla 6.16 Predicciones con la red FBR**

Corrida	Peso ( $y_1$ )	SPC 2 ( $y_2$ )	SPC 3 ( $y_3$ )
1	0.909	0.46	0.85
2	0.915	0.76	0.90
3	0.911	0.40	0.83
4	0.915	0.47	0.83
<b>5</b>	<b>0.915</b>	<b>0.13</b>	<b>0.75</b>
6	0.911	0.18	0.78
7	0.912	0.34	0.81
8	0.913	0.32	0.80
9	0.911	0.39	0.83
10	0.915	0.48	0.83
11	0.878	1.18	1.16
12	0.911	0.40	0.83
13	0.914	0.47	0.83
<b>14</b>	<b>0.933</b>	<b>-0.62</b>	<b>0.49</b>
15	0.912	-0.00054	0.73
<b>16</b>	<b>0.912</b>	<b>-0.00054</b>	<b>0.73</b>
17	0.912	-0.00054	0.73
18	0.912	-0.00054	0.73
19	0.912	-0.00054	0.73
20	0.912	-0.00054	0.73

Para la validación en campo se utilizaron los datos y parámetros de la corrida 14, 16 y 5 (ver Tabla 6.16) ya que son los valores con mejores resultados en las respuestas del proceso.

Los parámetros para estas corridas fueron las siguientes: para la primera prueba se utilizó la corrida 14 con la Temperatura del metal a 405°C,

Temperatura del molde de 230°C y 34.09 segundos de Basculamiento. El resultado de esta primera prueba fue de 0.934 kg de peso, -0.22 mm de SPC 2 y 0.17 mm de SPC 3, contra una predicción de 0.933 kg de peso, -0.62 mm de SPC 2 y 0.49 mm de SPC 3.

Para la segunda prueba se utilizó la corrida 16, la temperatura del metal fue de 405°C, Temperatura del molde de 230°C y 24 segundos de Basculamiento. El resultado de la segunda prueba fue de 0.908 kg de peso, 0 mm de SPC 2 y 0.7 mm de SPC 3, contra una predicción de 0.912 kg en peso, -0.00054 mm en SPC 2 y 0.7383 mm en SPC 3.

La tercera y última prueba consistió en la corrida 5, 390°C de Temperatura del metal, 210°C de Temperatura del molde y 30 segundos de Basculamiento. Los resultados fueron 0.917 kg de Peso, -0.16 mm en SPC 2 y 0.04 mm en SPC 3, contra una predicción de 0.915 kg de Peso, 0.13 mm en SPC 2 y 0.75 mm en SPC 3.

En la Tabla 6.17 se muestran los intervalos de confianza para los pesos y el bias, estimados mediante la ecuación (4.53).

**Tabla 6.17 Intervalos de confianza para los pesos**

	Límite inferior	$w_{j1}$	Límite superior	Límite inferior	$w_{j2}$	Límite superior	Límite inferior	$w_{j3}$	Límite superior
Bias	0.8685	0.9122	0.9559	-1.3739	0.5693	2.5126	-0.1891	0.8453	1.8797
$w_1$	-0.0444	-0.0007	0.0430	-3.0228	-1.0796	0.8637	-0.5238	0.5106	1.5449
$w_2$	-0.0995	-0.0558	-0.0121	0.6832	2.6264	4.5697	0.0726	1.1069	2.1413
$w_3$	0.0219	0.0656	0.1093	-6.6060	-4.6627	-2.7195	-2.9094	-1.8751	-0.8407

En la Tabla 6.18 se muestran los intervalos de confianza para la respuesta  $y_1$ : *peso*,  $y_2$ : *SPC 2* y  $y_3$ : *SPC 3* respectivamente, calculados mediante la ecuación (4.54).

**Tabla 6.18 Intervalos de confianza para las predicciones**

<b>Límite inferior</b>	$\hat{y}_0$ : <i>Peso</i>	<b>Límite superior</b>	<b>Límite inferior</b>	$\hat{y}_0$ : <i>SPC 2</i>	<b>Límite superior</b>	<b>Límite inferior</b>	$\hat{y}_0$ : <i>SPC 3</i>	<b>Límite superior</b>
0.8396	0.909	0.9697	-1.93	0.46	3.78	-0.48	0.85	2.47
0.842	0.915	0.972	-2.05	0.76	3.66	-0.52	0.9	2.42
0.8473	0.911	0.9773	-2.3	0.4	3.42	-0.63	0.83	2.32
0.8472	0.915	0.9772	-2.29	0.47	3.42	-0.63	0.83	2.32
0.8469	0.915	0.977	-2.31	0.13	3.41	-0.63	0.75	2.31
0.8464	0.911	0.9764	-2.78	0.18	2.93	-0.39	0.78	2.56
0.8611	0.912	0.9911	-3.27	0.34	2.44	-1.02	0.81	1.92
0.8535	0.913	0.9835	-2.84	0.32	2.87	-0.76	0.8	2.19
0.8474	0.911	0.9775	-2.32	0.39	3.39	-0.64	0.83	2.31
0.847	0.915	0.977	-2.47	0.48	3.24	-0.54	0.83	2.41
0.8329	0.878	0.963	-1.62	1.18	4.09	-0.34	1.16	2.61
0.8506	0.911	0.9806	-2.53	0.4	3.18	-0.72	0.83	2.22
0.8454	0.914	0.9755	-2.2	0.47	3.51	-0.59	0.83	2.35
0.8617	0.933	0.9917	-3.42	-0.62	2.29	-1	0.49	1.95
0.8464	0.912	0.9765	-2.9	-0.00054	2.82	-0.76	0.73	2.19
0.8464	0.912	0.9765	-2.9	-0.00054	2.82	-0.76	0.73	2.19
0.8464	0.912	0.9765	-2.9	-0.00054	2.82	-0.76	0.73	2.19
0.8464	0.912	0.9765	-2.9	-0.00054	2.82	-0.76	0.73	2.19
0.8464	0.912	0.9765	-2.9	-0.00054	2.82	-0.76	0.73	2.19
0.8464	0.912	0.9765	-2.9	-0.00054	2.82	-0.76	0.73	2.19

Las Figuras 6.6, 6.7 y 6.8 muestran los intervalos de confianza para las predicciones y las réplicas utilizadas para la validación en campo. En dichas figuras se puede comprobar que las réplicas caen dentro de los intervalos de confianza.

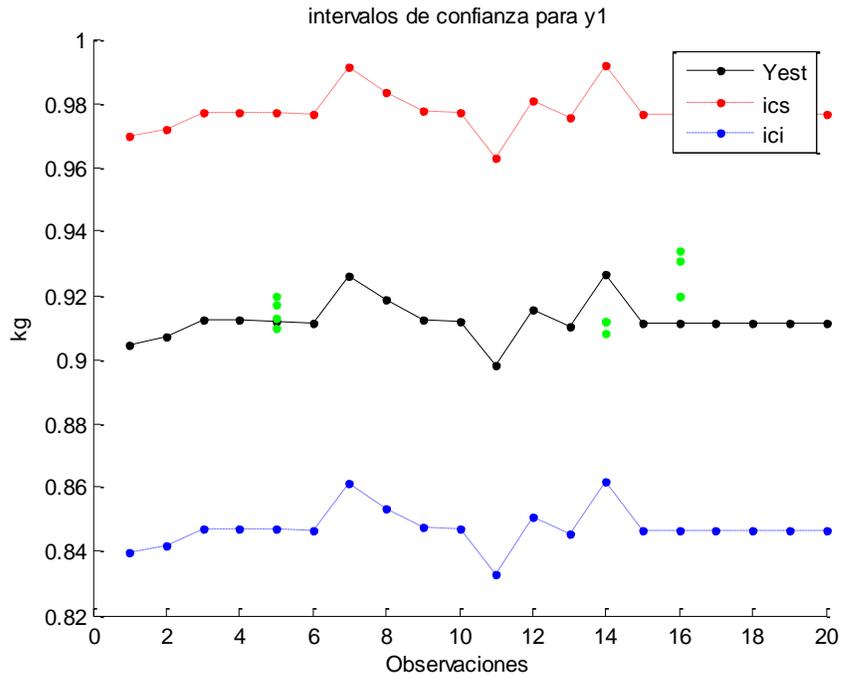


Figura 6.6 Intervalos de confianza para Peso

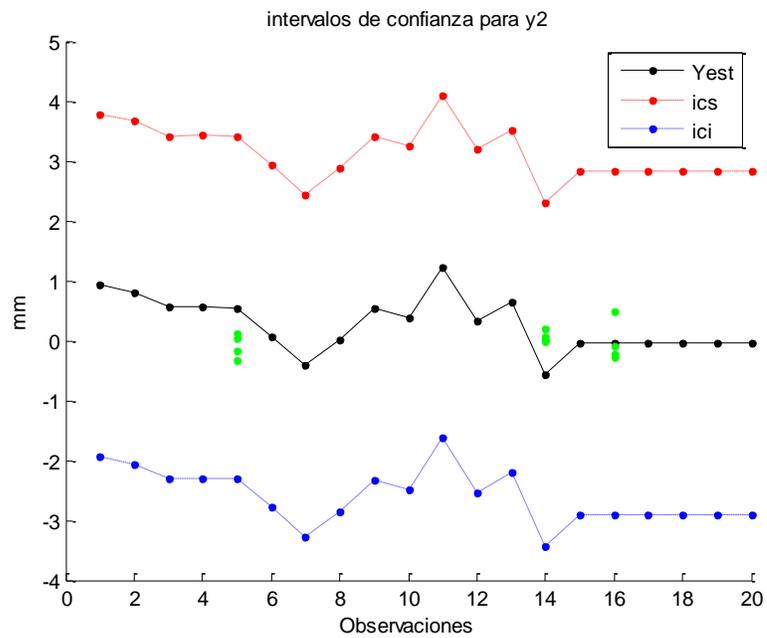
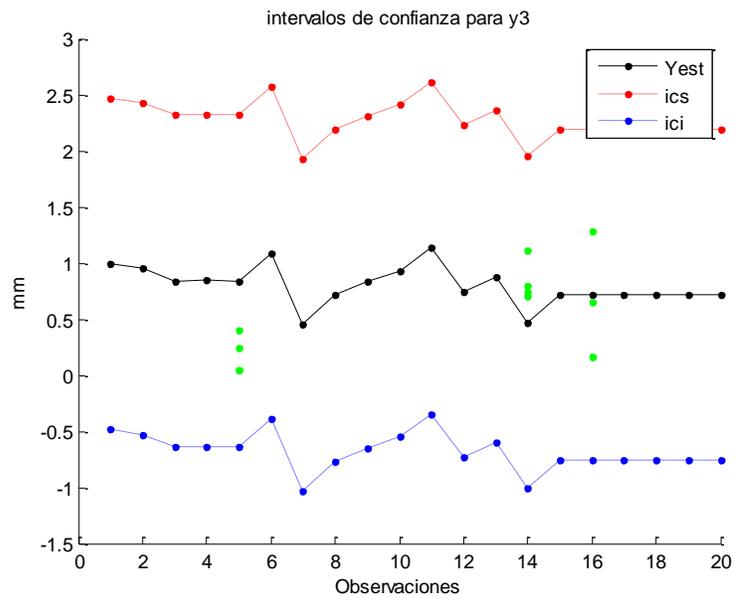


Figura 6.7 Intervalos de confianza para SPC 2



**Figura 6.8 Intervalos de confianza para SPC 3**

Así, es posible observar que la red neuronal de Función de Base Radial generalizada realiza una predicción adecuada ya que las pruebas realizadas encuentran resultados precisos en el proceso de fundición en molde permanente y por gravedad.

## Capítulo 7

### Conclusiones

Debido a los rápidos cambios en los métodos de diseño y la demanda de productos de calidad, las empresas de manufactura han visto incrementada la dificultad en la toma de decisiones en sus procesos (Ahilan et al. 2013). Bajo estas circunstancias, es común encontrar casos donde a prueba y error tratan de ajustar parámetros y variables, pero dicha práctica genera gastos en diversos sentidos (Benyounis y Olabi, 2008). El uso de herramientas numéricas elimina los gastos del uso de estas prácticas, y ayudan a mejorar y robustecer los procesos de manufactura. La red neuronal de Función de Base Radial es una herramienta con un uso importante en la modelación de procesos de manufactura, sin embargo, en la literatura existen modelos que analizan procesos por medio de esta red donde solo se modela una salida, y dadas las características de los procesos no siempre es conveniente modelar varias respuestas de manera independiente.

Debido a estas dificultades se investigó y desarrolló el presente trabajo, donde se presenta la modelación mediante una red neuronal de Función de Base Radial y su análisis mediante técnicas estadísticas como el ANOVA, MANOVA y métricas de validación ( $R^2$ ,  $\eta^2$ , etc.), verificando antes si se cumplen los supuestos para los análisis antes mencionados. Lo anterior como una

metodología general propuesta para la predicción de las características de calidad deseadas propiciando procesos más robustos.

Para cumplir con **el objetivo general** de esta investigación, el cual es: desarrollar un modelo generalizado de red neuronal de Función de Base Radial que cumpla con los supuestos para el análisis de varianza multivariado, se desarrolló un modelo generalizado de función de base radial capaz de predecir procesos con una o varias respuestas, cumpliendo también el **objetivo específico uno**. Para el cumplimiento de los supuestos para el análisis de varianza univariado y multivariado, dichos supuestos son; para el análisis univariado: normalidad, independencia y homocedasticidad, y pueden ser verificados mediante las pruebas de Anderson-Darling, Durbin-Watson y White respectivamente. Para el análisis multivariado los supuestos son: normalidad, linealidad, distribución independiente y homocedasticidad, estos últimos se verifican mediante las pruebas de Henze-Zirckler, Shapiro-Wilk y Mardia, coeficiente de correlación de Pearson, Durbin-Watson y White respectivamente. Con el cumplimiento de estos supuestos se puede asegurar que el modelo representa los procesos de manera correcta y las decisiones que se tomen en base a estos serán más acertadas, cumpliendo **el objetivo específico dos**.

La modelación univariada o multivariada se determinó mediante la relación entre variables de salida. Después, se aplicó el algoritmo de agrupamiento por algoritmo genético para la construcción de la matriz de centros y se utilizó como función de evaluación el coeficiente de determinación  $R^2$  para mejorar la precisión de la predicción de la red. En ambos casos, el objetivo es minimizar las distancias entre los centros y los patrones de entrada, al maximizar dichos coeficientes se garantiza una mejor predicción y no genera la inestabilidad que otros métodos comúnmente utilizados en la red de base radial generan como por ejemplo k-medias. La diferencia está en la forma de calcular los coeficientes, utilizando la extensión multivariada del coeficiente de

determinación  $R_q^2$ . Con lo anterior se respondieron **las preguntas de investigación** 3.- ¿cómo será la estimación de los centros de la red en el caso multivariado? y 4.- ¿cómo es la estimación de los pesos en las redes que consideran varias respuestas? También se cumplieron **los objetivos específicos** 3.- obtener los centros para la función de base radial sin generar inestabilidad en la predicción de la red, y 4.- estimar los pesos de la red mediante técnicas estadísticas multivariadas.

Después de determinar los centros, se mide la distancia entre estos y los patrones de entrada mediante la distancia de Mahalanobis, para construir la matriz de interpolación  $G$ . Dicha matriz de interpolación depende de la función de base radial, de las cuales se planteó que existen muchas. En DeLeón-Delgado et al., (2015) se propuso utilizar los resultados del algoritmo genético como criterio de selección de la función de base radial indicada para el proceso, de manera que la función que minimice las distancias entre los centros y los patrones de entrada obtendrá el máximo en la función objetivo y se seleccionará esa función de base radial. Con esto se cumplió con **el objetivo específico 5**.- Analizar los efectos de diferentes funciones de base radial en la red neuronal multivariada. Por otra parte, los pesos sinápticos son calculados por medio de mínimos cuadrados, que al multiplicarse por la matriz  $G$  se obtienen las salidas estimadas.

Como parte de la presente investigación, se analizó un proceso de fundición en molde permanente. Por medio de un algoritmo genético se estimaron los centros de la red, de manera que el algoritmo encontró un óptimo global al minimizar las distancias entre los centros generados y los patrones de entrada. Esto se comprobó mediante el estadístico  $R_q^2$ , el cual fue maximizado al cumplir esta condición. Para los pesos sinápticos se utilizó el estimador de mínimos cuadrados multivariados. Cumpliendo **el objetivo específico 6**; la aplicación de la metodología en campo. El molde de proceso cuenta con seis

cavidades de las cuales cada cavidad es una pieza. Para seleccionar qué cavidad analizar se utilizó un ANOVA de un factor, de donde se concluye que las cavidades con mayor variación del proceso son la P para la respuesta  $y_1$ : peso, Q para  $y_2$ : SPC 2, y Q para  $y_3$ : SPC 3. El objetivo es reducir la variación del proceso por lo que se seleccionaron esas cavidades para el análisis. El modelo de red neuronal fue multivariado debido a la correlación entre las respuestas, y al cumplir con los supuestos estadísticos para el análisis multivariado de varianza, se verificó que el modelo explica de manera correcta la variabilidad del proceso y que **la variable que más impacta en esta variabilidad es la temperatura del molde.**

Al obtener los resultados se verificaron los supuestos para el análisis de varianza multivariado. La clave está en determinar si las variables dependientes están correlacionadas, de esta forma el supuesto de linealidad entre variables dependientes para el análisis de varianza multivariado se cumple. Al cumplirse esté supuesto y el de normalidad puede realizarse dicho análisis. Las ventajas de realizar este análisis en lugar de un ANOVA para cada variable dependiente es que se trabaja con todas las variables simultáneamente, y se reduce la inflación del error tipo I presentado al ignorar las relaciones entre las variables.

Como trabajo futuro se propone el uso de otros algoritmos evolutivos en la red para la estimación de los centros, por ejemplo, el algoritmo por enjambre de partículas (PSO). Además de la optimización de procesos con varias respuestas, verificar si esta optimización será multi-objetivo.

Además, el trabajo futuro contemplaría temas como la selección de la función de base radial que mejor se adapte al comportamiento del proceso y cómo regresar los datos interpolados por la función de base radial a su dimensión original.

## Bibliografía

Ahilan, C., Kumanan, S., Sivakumaran, N., Edwin, J. (2013). *Modeling and prediction of machining quality in CNC turning process using intelligent hybrid decision making tools*. Applied soft computing 13 pp. 1543-1551.

Akhtar, T., Shoemaker, C. (2015). *Multi objective optimization of computationally expensive multi-modal functions with RBF surrogates and multi-rule selection*. Global Optimization 64 pp. 17-32.

Andrews, D., Gnanadesikan, R., Warner, J. (1971). *Transformations of multivariate data*. Biometrics 27 pp. 825-840.

Bashiri, M., Hasan, M. (2014). *An optimality probability index for correlated multiple responses*. Communications in statistics-Theory and methods 43 pp. 4324-4336.

Benyounis, K., Olabi, A. (2008). *Optimization of different welding process using statistical and numerical approaches – A reference guide*. Advances in engineering-Software 39 pp. 483-496.

Broomhead D., Lowe D. S. (1988). *Multi-variable functional interpolation and adaptive networks*. Complex Systems 2 pp. 321-355.

- Cevallos J. (2008). *Redes Neuronales de Base Radial aplicadas a la mejora de la calidad*. Industrial Data 11-2 pp. 63-72.
- Cervantes, R., Arteaga, R., Vázquez, M., Ojeda, W. (2012). *Red neuronal artificial base radial en la estimación de la evapotranspiración de referencia*. Revista mexicana de Ciencias Agrícolas 14 pp. 698-704.
- Chandrasekaran, M., Muralidhar, M., Murali, C., Dixit, U. (2010). *Application of soft computing techniques in machining performance prediction and optimization: a literatura review*. International Journal of Advanced Manufacturing Technology 46 pp. 445-464.
- Chen, S., Billings, S., Luo, W. (1989). *Orthogonal least squares methods and their application to non-linear system identification*. International Journal of Control 50 pp. 1873-1896.
- Cover, T. (1965). *Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition*. IEEE Transactions on Electronic Computers EC-14 pp. 326-334.
- DeLeón-Delgado, H., Praga-Alejo, R., González-González, D. (2015). *Different radial basis functions in a welding process*. Proceedings of the 4th annual world conference of the Society for Industrial and Systems Engineering, Fort Lauderdale, FL. USA.
- De Paiva, A., Gomes, J., Santana, R., Coradi, R., Balestrassi, P. (2014). *A multivariate robust parameter optimization approach based on principal component analysis with combined arrays*. Computers & Industrial Engineering 74 pp. 186-198.

- Dieter, G. (1988). *Mechanical Metallurgy*, Third edition. Mc Graw Hill.
- Du D., Li K., Fei M. (2010). *A fast-multi-output RBF neural network construction method*. Neurocomputing 73 pp. 2196-2202.
- Gomes, C., Canedo, J. (2015). *Neural network of Gaussian radial basis function applied to the problem of identification of nuclear accidents in a PWR nuclear power plant*. Annals of Nuclear Energy 77 pp.285-293.
- Gomes, J., Paiva, A., Costa, S., Balestrassi, P., Paiva, E. (2013). *Weighted multivariate mean square error for processes optimization: A case study on flux-cored arc welding for stainless steel claddings*. European journal of operational research 226 pp. 522-535.
- González-González, D., Praga-Alejo, R., Mendoza-Santillán, G., Piña-Monarez, M. (2011). *Modeling a process with several correlated responses*. Proceeding of the 16th annual International Conference on Industrial Engineering theory, applications, and practice, Stuttgart, Germany.
- Haykin, S. (1999). *Neural Networks a Comprehensive Foundation*. Pearson Prentice Hall, Delhi, India.
- Haji, K., Pirmoradi, Z., Cheng, G., Wang, G. (2015). *Decomposition for large-scale global optimization based on quantified variable correlations uncovered by metamodelling*. Engineering Optimization 47-4 pp. 429-452.
- He, H., Ji-cheng, B., Ze-sheng, Lu., Yong-feng, G. (2009). *Electrode wear prediction in milling electrical discharge machining based on radial basis function neural network*. Shanghai Jiaotong 14(6).

- Henze, N., Zirkler, B. (1990). *A class of invariant consistent test for multivariate normality*. Communications in statistics – Theory and methods 19 pp. 3595-3617.
- Holland, J. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press: Ann Arbor, USA.
- Huda, S., Abdollahian, M., Mammadov, M., Yearwood, J., Ahmed, S., Sultan, I. (2014). *A hybrid wrapper-filter approach to detect the source (s) of out-of-control signals in multivariate manufacturing process*. European Journal of Operational Research 237 pp. 857-870.
- Jiang, B., Zhang, F., Sun, Y., Zhou, X., Jiabin, Dong., Luhong, Z. (2014). *Modeling and optimization for curing of polymer flooding using an artificial neural network and a genetic algorithm*. Journal of the Taiwan Institute of Chemical Engineers 45 pp. 2217-2224.
- Jie, H., Wu, Y., Ding, J. (2014). *An adaptive metamodel-based global optimization algorithm for black-box type problems*. Engineering optimization 47 pp. 1459-1480.
- Karimi, F., Rafiee, S., Taheri, A., Karimi, M. (2012). *Optimization of an air-drying process for Artemisia absinthium leaves using response surface and artificial neural network models*. Journal of the Taiwan Institute of Chemical Engineers 43 pp 29-39.
- Lee, S., Sung, W. (2013). *A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis*. Expert Systems with applications 40 pp. 2941-2946.

- Liu, H., Xu, S., Wang, X., Wu, J., Sond, Y. (2014). *A global optimization algorithm for simulation-based problems via the extended DIRECT scheme*. Engineering Optimization 47 pp. 1441-1458.
- Lopes, L., Gomes, J., de Paiva, A., Barca, L., Ferreira, J., Balestrassi, P. (2013). *A multivariate surface roughness modeling and optimization under conditions of uncertainty*. Measurement 46 pp. 2555-2568.
- Mardia, K. (1970). *Measures of multivariate skewness and kurtosis with applications*. Biometrika 57 pp. 519-530.
- Mardia, K. (1980). *Tests of univariate and multivariate normality*. Handbook of statistics 1 pp. 279-320, North-Holland Company, Amsterdam.
- Mateo, J., Rieta, J. (2013). *Radial Basis Function Networks Applied to Efficient QRST Cancellation in Atrial Fibrillation*. Computers in Biology and Medicine 43 pp. 154-163.
- McCulloch, W., Pitts, W. (1943). *A logical calculus of the ideas immanent in nervous activity*. Bulletin of mathematical biophysics 5 pp. 115–133.
- Mecklin, C., Mundfrom, D. (2004). *An appraisal and bibliography of tests for multivariate normality*. International Statistical Review 72 pp. 123-138.
- Micchelli, C. (1986). *Interpolation of scattered data: distance matrices and conditionally positive definite functions*. Constructive approximation 2 pp. 11-22.

- Mollah, A., Pratihari, D. (2008). *Modeling of TIG welding and abrasive flow machining processes using radial basis function networks*. Int. J. Adv. Manuf. Technol 37 pp. 937-952.
- Montgomery, D. C., Peck, E. A., Vining, G. G. (2006). *Introducción al análisis de Regresión Lineal*. Continental, México.
- Montgomery, D. C. (2006). *Diseño y análisis de experimentos*. Limusa, México.
- Nelles, O. (2001). *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*. Springer-Verlag Berlin Heidelberg, Germany.
- Patel, M., Krishna, P., Parappagoudar, M. (2016). *Squeeze casting process modeling by a conventional statistical regression analysis approach*. Applied mathematical modeling 40 pp. 6869–6888.
- Pérez, M., Rivera, A., Carmona, C., del Jesús, M. (2014). *Training algorithms for radial basis function networks to tackle learning processes with imbalanced data-sets*. Applied Soft Computing 25 pp. 26-39.
- Powell, M. (1987). *Radial basis function approximations to polynomials*. Dundee niennial Numerical Analysis Conference.
- Praga, R., González, D., Cantú, M., Perez, P., Torres, L., Flores, B. (2013). *Statistical inference in a redesigned Radial Basis Function neural network*. Engineering applications of artificial intelligence 26 pp. 1881-1891.

- Qasem, S., Shamsuddin, S. (2011). *Radial basis function network based on time variant multi-objective particle swarm optimization for medical diseases diagnosis*. Applied soft computing 11 pp. 1427–1438.
- Rajendra, M., Shankar, M. (2015). *Improved complex-valued Radial Basis Function (ICRBF) neural networks on multiple crack identification*. Applied Soft Computing 28 pp. 285-300.
- Rencher, A. (2002). *Methods of multivariate analysis*. Wiley-Interscience, USA.
- Romeu, J., Ozturk, A. (1993). *A comparative study of goodness-of-fit tests for multivariate normality*. Journal of Multivariate Analysis 46 pp. 309-334.
- Sadeghzadeh, J., Sadeghi, M., Zarrinpashne, S., Irandoukht, A. (2013). *Optimization of OCM reactions over Na-W-Mn/SiO<sub>2</sub> catalyst at elevated pressure using artificial neural network and response surface methodology*. Scientia Iranica C 20(3) pp. 617-625.
- Salmasnia, A., Baradaran, R., Mohajer, M. (2012). *A novel approach for optimization of correlated multiple responses based on desirability function and fuzzy logics*. Neurocomputing 91 pp. 56-66.
- Sampaio, D., Vasconcelos, C., Simões, S., Antonio, V. (2005). *Comparison between genetic algorithms and response surface methodology in GMAW welding optimization*. Journal of Materials Processing Technology 160 pp. 70-76.
- Schafföner, S., Freitag, L., Hubálková, J., Aneziris, C. (2016). *Functional composites based on refractories produced by pressure slip casting*. Journal of the European Ceramic Society 36 pp. 2109–2117.

- Shahlaei, M., Bahrami, G., Abdolmaleki, S., Sadrjavadi, K., Bagher, M. (2015). *Application of unfolded principal component analysis-radial basis function neural network for determination of celecoxib in human serum by three-dimensional excitation-emission matrix fluorescence*. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 138 pp. 675-683.
- Timelli, G., Caliarì, D., Rakhmonov, J. (2016). *Influence of process parameters and Sr addition on the microstructure and casting defects of LPDC A356 alloy for engine blocks*. Journal of materials science & technology 32 pp. 515–523.
- Timm, N. (2002). *Applied Multivariate Analysis*. Springer Texts in Statistics, USA.
- Tyan, M., Van Nguyen, N., Lee, J. (2014). *Improving variable-fidelity modelling by exploring global design space and radial basis function networks for aerofoil design*. Engineering Optimization 47 pp. 885-908.
- Ugrasen, G., Ravindra, H.V., Naveen Prakash, G.V. y Keshavamurthy R. (2014). *Process optimization and estimation of machining performances using artificial neural network in wire EDM*. Procedia Materials Science 6 pp. 1752 – 1760.
- White, H. (1980). *A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity*. Econometrica 48 pp. 817-838.
- Yen, J., Langari, R. (2000). *Fuzzy logic: intelligence, control, and information*. University of Michigan: Prentice Hall, USA.