

**CUANTIFICACION DE PROPIEDADES MECANICAS Y QUIMICAS DEL
POLIPROPILENO USANDO ESPECTROSCOPIA DE RESONANCIA
MAGNÉTICA NUCLEAR PROTÓNICA (^1H -RMN).**

ALVEIRO HERNANDEZ MORELO



**UNIVERSIDAD DE CARTAGENA
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
PROGRAMA DE MESTRÍA EN QUÍMICA**

**CARTAGENA DE INDIAS D.T. Y C.,
MARZO 2015**

**CUANTIFICACION DE PROPIEDADES MECANICAS Y QUIMICAS DEL
POLIPROPILENO USANDO ESPECTROSCOPIA DE RESONANCIA
MAGNÉTICA NUCLEAR PROTÓNICA (¹H-RMN).**

ALVEIRO HERNANDEZ MORELO

Trabajo de grado presentado como requisito parcial para optar el título de
Magíster en Química

BORIS JOHNSON RESTREPO Q.F., *Ph.D.*

Director del trabajo



**UNIVERSIDAD DE CARTAGENA
FACULTAD DE CIENCIAS EXACTAS Y NATURALES
PROGRAMA DE MESTRÍA EN QUÍMICA
CARTAGENA DE INDIAS D.T. Y C.,
MARZO 2015**

AGRADECIMIENTOS

Durante el desarrollo de mis estudios de Maestría en Química en la Universidad de Cartagena, recibí la colaboración directa e indirecta de diferentes personas e instituciones, a las cuales quiero manifestarles mis sinceros agradecimientos: al grupo de Investigación de Química y Medio Ambiente liderado por el Profesor Boris Johnson Restrepo, Ph.D. como director del trabajo de grado. A Jorge Mario Roperó Vega, Ph.D y Harold Gómez Estrada, Ph.D. por sus invaluable aportes y soporte científico como evaluadores del presente trabajo de investigación, que sirvieron como una guía permanente. A los químicos de la Universidad de Cartagena y miembros del grupo de Investigación de Química y Medio Ambiente, Julio Manuel Banquet Teran y Yeison Enrique Cuadrado Noriega por su tiempo y disponibilidad para ayudar y asesorar durante el desarrollo de esta investigación.

También quiero expresar mi gratitud a la empresa Polipropileno del Caribe S.A. (Essentia by Propilco) y en especial a Jaime Santoya Tello, Jefe del Laboratorio; Juan Diego Mejía Mejía, Vicepresidente de Abastecimiento y Logística; Lázaro Navarro Ruiz Gerente de Producción, por aprobar y apoyar la realización de este trabajo de investigación; y a todos los analistas del laboratorio de Propilco por su dedicación en la realización de los ensayos.

DEDICATORIA

En primer lugar a Dios, el todo poderoso, a quién le agradezco por mi existencia, mi fe, salud, constancia y conocimientos que me permitieron alcanzar este triunfo. Infinitas gracias Dios.

A mi madre Isabel que es la inspiración y el motor que ha impulsado todos mis triunfos. A mi padre Gregorio por su sabiduría de la vida. Los Amo mucho.

A mi querida familia, mi esposa Albertina, mis hijos Luis Eduardo, Alveiro Jr., y María Camila, por su amor incondicional, quienes me han acompañado en los malos y buenos momentos. Son el centro de mí vivir. Los adoro.

A mi abuelita Inés, quien con su inmensa sabiduría me enseñó a buscar ser mejor cada día. El señor la tenga en su gloria.

A todos mis familiares y amigos que contribuyeron directa e indirectamente en alcanzar este gran triunfo en el nombre de Dios. Los llevo en mi mente y en mi corazón.

Nota de aceptación

Presidente del jurado

Jurado

Jurado

Cartagena de Indias, Marzo 2015

TABLA DE CONTENIDO

1. INTRODUCCIÓN	15
2. MARCO TEÓRICO	17
2.1 POLIPROPILENO	17
2.2 COPOLIMERIZACIÓN DE MOLÉCULAS DE POLIPROPILENO	21
2.3 PROPIEDADES MECÁNICAS DEL POLIPROPILENO.....	28
2.4 PROCESOS DE TRANSFORMACIÓN	28
2.5 PROPIEDADES MECÁNICAS	31
2.5.1 Resistencia a la tensión a la cedencia (Tensile strength at yield)	34
2.5.2 El módulo de flexión 1% secante	35
2.6 ANÁLISIS EN MUESTRAS SOLIDAS DE POLIPROPILENO USANDO CURVAS DE DECAIMIENTO MAGNÉTICO INDUCIDO (FID) POR RESONANCIA MAGNÉTICA NUCLEAR PROTÓNICA (¹ H-RMN) DE BAJA RESOLUCIÓN.	36
2.7 QUIMIOMETRÍA.....	46
2.7.1 Pre-tratamientos de los Datos	47
2.7.2 Análisis Multivariado de Datos	48
2.7.3 Modelos de Regresión Lineal Múltiple	71
3. MATERIALES Y MÉTODOS	86
3.1 MUESTRAS EN FORMA DE POLVO Y DE PELLETS.....	86
3.2 ADQUISICIÓN DE LAS CURVAS DE FID.....	87
3.3 DETERMINACIÓN DEL PORCENTAJE DE ETILENO USANDO INFRARROJO MEDIO.....	87
3.4 MEDICIÓN DEL MÓDULO DE FLEXIÓN 1 % SECANTE.....	88
3.5 PRE-PROCESAMIENTO DE LOS DATOS.....	89
3.6 GENERACIÓN Y EVALUACIÓN DE LOS MODELOS DE PREDICCIÓN OBTENIDOS	89
4. RESULTADOS.....	91
4.1 ESPECTROS ¹ H-RMN OBTENIDOS PARA MUESTRAS DE PP HOMOPOLÍMERO Y COPOLÍMERO	91

4.2 TRANSFORMACIONES QUIMIOMÉTRICAS APLICADAS A LOS ESPECTROS ^1H -RMN DE MUESTRAS DE PP.....	95
4.3 ANÁLISIS DE COMPONENTES PRINCIPALES DE LAS CURVAS FID/ ^1H -RMN	98
4.3.2 INTERPRETACIÓN DE LOS LOADINGS DEL PCA PARA LAS CURVAS FID/ ^1H -RMN	101
4.4 PREDICCIÓN DE PROPIEDADES FISCOQUÍMICAS EN MUESTRAS DE POLIPROPILENO USANDO ESPECTROSCOPIA DE ^1H -RMN	106
4.4.1 Generación y selección de descriptores matemáticos para curvas FID ^1H -RMN	106
4.4.2 ANÁLISIS DE COMPONENTE PRINCIPAL APLICADO A LOS DESCRIPTORES MATEMÁTICOS DE CURVAS FID/ ^1H -RMN	109
4.4.3 INTERPRETACIÓN DE LOS PC LOADINGS PARA DESCRIPTORES MATEMÁTICOS DE CURVAS FID/ ^1H -RMN	110
4.5 CALIBRACIÓN MULTIVARIADA CON REGRESIÓN POR MÍNIMOS CUADRADOS PARCIALES (PLS).....	112
4.6 PREDICCIÓN DE LA CONCENTRACIÓN DE ETILENO EN MUESTRAS DE CEP-LER USANDO ESPECTROSCOPIA ^1H -RMN	112
4.7 CALCULO DE LA EXACTITUD DEL MODELO PLS PARA PREDICCIÓN DE ETILENO	116
4.8 PREDICCIÓN DE LA RESISTENCIA A LA TENSIÓN PARA MUESTRAS DE PP-H, CEP-LER, CEP-R Y CEP-I USANDO CURVAS DE FID/ ^1H -RMN	118
4.9 EXACTITUD DE MODELOS DE PREDICCIÓN DE TENSILE PARA LAS DIFERENTES FAMILIAS DE PP	121
4.10 MODELO DE PREDICCIÓN DEL MÓDULO DE FLEXIÓN 1% SECANTE EN MUESTRAS DE POLIPROPILENO USANDO ESPECTROSCOPIA ^1H -RMN.....	123
4.11 EXACTITUD DEL MODELO DE PREDICCIÓN DEL MODELO DE FLEXIÓN 1% SECANTE PARA MUESTRAS DE POLIPROPILENO	124
5. CONCLUSIONES	127
6. BIBLIOGRAFÍA	128

LISTA DE ABREVIATURAS

ASTM:	American Standard for Testing and Materials.
BOPP:	Película de Polipropileno Biaxialmente Orientada (Bi-axially Oriented Polypropylene).
CPE-i:	Polipropileno copolímero de impacto propileno/etileno.
CPE-LER:	Polipropileno copolímero random polipropileno/etileno de bajo contenido de etileno.
CPE-r:	Polipropileno copolímero random propileno/etileno.
CV:	Validación Cruzada (Cross Validation).
DSC:	Calorímetro Diferencial de Barrido (Differential Scanning Calorimetry).
Ec:	Porcentaje en peso de etileno en copolímero de propileno-etileno o fase caucho.
Et:	Porcentaje total de etileno en peso en la cadena polimérica.
Fc:	Porcentaje de fase copolímero o caucho en el polímero total.
FT:	Transformación de Fourier.
GLS:	Mínimos Cuadrados Generalizados (Generalized Least Squares).
¹H-RMN:	Resonancia Magnética Nuclear Protónica.
IR:	Infrarrojo (Infra Red).
ISO:	International Standard Organization.
KNN:	Algoritmo de k-Vecinos más cercanos (k-nearest neighbours algorithm).
LDA:	Análisis Discriminante Lineal (Lineal Discriminant Analysis).
MFR:	Rata de flujo de fundido (Melt Flow Rate).
MID:	Infrarrojo Medio (Mid InfraRed).
MLR:	Regresión Lineal Múltiple (Multiple Linear Regression).
MMC:	Método de Mínimos Cuadrados.
MSC:	Corrección de Dispersión Multiplicativa (MultiplicativeScatterCorrection).
NIPALS:	Estimación de los parámetros de un modelo no lineal con regresiones simples entre los datos una parte de los parámetros (Non-Linear estimation by Iterative Partial Least Square).
NIR:	Infrarrojo Cercano(Near InfraRed).
PCA:	Análisis de Componente Principal (Principal Component Analysis).
PCR:	Regresión por Componentes Principales (Principal Component Regression).
PE:	Polietileno.
PLSR:	Regresión Parcial por Mínimos Cuadrados (Partial Least Squares Regression).
PP:	Polipropileno.
PP-a:	Polipropileno atáctico.
PP-h:	Polipropileno homopolímero.
PP-i:	Polipropileno isotáctico.

PP-s:	Polipropileno sindiotáctico.
PRESS:	Suma Residual Predictiva de Cuadrados (Predicted Residual Sums of Squares).
PSI:	Libras-Fuerza por Pulgada Cuadrada (Pounds per Squarde Inch).
QDA:	Análisis Discriminante Cuadrático (Quadratic Discriminant Analysis).
RFP:	Pulso de Radio Frecuencia.
RMSE:	Error Cuadrático Medio (Root Mean Square Error).
RMSEC:	Error Cuadrático Medio de Calibración (Root Mean Square Error of Calibration).
RMSEP:	Error Cuadrático Medio de Predicción (Root Mean Square Error of Prediction).
S:	Similitud.
SCE:	Suma de Cuadrados del Error.
SCR:	Suma de Cuadrados Regresionales.
SIMCA:	Modelo Suave e Independiente de Analogías de Clases (Soft Independent Modeling of Class Analogy).
SNR:	Relación Señal-Ruido.
SNV:	Variación Normal Estandarizada (Standard Normal Variate)
SVD:	Descomposición en Valores Singulares (Singular Value Decomposition).
Tg:	Temperatura de Transición Vítreá.
TWQ:	Película Tubular enfriada por baño de agua (Tubular Water Quenched Film).

LISTA DE FIGURAS

	Pág.
Figura 1. Formulas estructurales a) representación del monómero de propileno, b) representación del polipropileno.....	18
Figura 2. Esquema que muestra el ciclo de polimerización del polipropileno.	18
Figura 3. Configuración estereoquímica del PP a) configuración isotáctica b) configuración sindiotáctica c) configuración atáctica.....	20
Figura 4. Copolímero random de PP. P = Propileno, E = Etileno.	23
Figura 5. Peso molecular, representado indirectamente por la rata de flujo de fundido - MFR (ASTM D 1238), sobre la resistencia al impacto Izod (ASTM D 256). Fuente:(Edward P. Moore, 1996).	26
Figura 6. Gráfica esfuerzo-deformación característica de un espécimen normalizado de Polipropileno, generada en un ensayo para determinar las propiedades mecánicas. (Shah, 1984).....	32
Figura 7. Diferentes tipos de extensión a nivel molecular del polipropileno durante la deformación mecánica. Tomado con modificación de Shah,1984.	33
Figura 8. Espécimen de ensayo para la prueba de tensión y flexión.....	34
Figura 9. Esquema de las fuerzas actuando sobre la flexión de un espécimen de polipropileno. (Tomado con modificaciones de McGraw-Hill Book Company).	36
Figura 10. Representación de los vectores resultantes en un proceso de precesión magnética para spines expuestos a un campo magnético externo propios de un proceso de ^1H -RMN durante el análisis del polipropileno.	39
Figura 11. Representación gráfica de la curva de decaimiento FID y de las tres curvas que modelan el proceso de decaimiento magnético.	42
Figura 12. Matriz X de n filas y n columnas.....	49
Figura 13. Representación gráfica de las matrices envueltas en un análisis de componente principal.	52
Figura 14. (a) Curvas FID/ ^1H -RMN generadas para muestras de PP-h, y los diferentes tipos de copolímeros propileno/etileno en forma de pellets (b) Curvas FID/ ^1H -RMN generadas para muestras de CPE-LER comparando la presentación en forma de polvo y en forma de pellets.	92
Figura 15. Corrección de línea base aplicada a curvas FID ^1H -RMN de muestras de PP en forma de pellets.....	96
Figura 16. SNV aplicada a curvas FID ^1H -RMN de muestras de PP sólido.....	97
Figura 17. Primera derivada aplicada a curvas FID ^1H -RMN de muestras de PP solido.	97
Figura 18. Segunda derivada aplicada a curvas FID ^1H -RMN de muestras de PP solido.....	98

Figura 19. Representación gráfica de los escores obtenidos en el PCA para curvas FID de muestras de PP a) scores PC1 vs Score PC2 b) score PC1 vs score PC3.....	100
Figura 20. Representación gráfica de los loadings para el análisis PCA de curvas FID/ ¹ H-RMN para muestras de PP en forma de pellets a) representación grafica del loading PC1 b) representación grafica del loading PC2.....	103
Figura 21. Representación gráfica de los escores obtenidos en el PCA para curvas FID de muestras de CPE-LER en polvo y pellets.....	105
Figura 22. Representación gráfica de los loadings obtenidos en el PCA para curvas FID de muestras de CPE-LER en polvo y pellets a) Loading PC 1 b) loading PC 2.	105
Figura 23. a) Representación gráfica de los descriptores secundarios obtenidos de la combinación de descriptores primarios b) Representación gráfica de los descriptores secundarios luego de un escalado.	108
Figura 24. Representación gráfica del análisis de componente principal de los scores PC 1 vs PC 2.....	109
Figura 25. Representación gráfica de los loading generados por el PCA para descriptores matemáticos de curvas FID ¹ H-RMN de muestras de PP. a) PC loading 1, b) PC loading 2, c) PC loading 3.	111
Figura 26. Predicción del modulo del contenido de etileno en muestras de polipropileno copolímero LER usando espectroscopia ¹ H-RMN y comparando con lecturas por Infrarrojo medio de las mismas muestras diferentes a las usadas como calibración y validación.....	117
Figura 27. Correlación del % etileno en CEP-LER del método de referencia IR medio y del método PLS para FID/ ¹ H-RMN	118
Figura 28. Valores de tensión medidos por el método de referencia y los valores predichos por el modelo PLS para las muestras de validación.	120
Figura 29. Comportamiento del valor medido y de la predicción del la resistencia a la tensión Secante para PP-h, copolímero LER, Random y de impacto usando espectroscopia ¹ H-RMN.	122
Figura 30. Modelo PLS log 10, empleando 2 factores en el modelo de predicción del módulo de flexión en pellets de PP.....	126
Figura 31. Comportamiento del valor medido y de la predicción del módulo de flexión 1% Secante para PP-h, copolímero LER, Random y de impacto usando espectroscopia ¹ H-RMN.	126

LISTA DE TABLAS

	Pág.
Tabla 1. Normas ASTM para determinar propiedades físico-químicas y mecánicas en el polipropileno.....	21
Tabla 2. Propiedades mecánicas típicas de las diferentes amilias de polipropileno	24
Tabla 3. Efecto de la naturaleza del PP y de la presencia de agente clarificante sobre la propiedad óptica opacidad.	28
Tabla 4. Algunos procesos de transformación y aplicaciones de uso final del Polipropileno.	30
Tabla 5. Esquema de tabla anova para una regresión lineal simple.....	68
Tabla 6. Esquema general de tabla anova de una regresión lineal simple	73
Tabla 7. Descriptores matemáticos primarios A , a , B , b , C y c para un grupo seleccionado de muestras estudiadas usando curvas FID generados por el software AzTec que controla y adquiere los datos del equipo de $^1\text{H-RMN}$ equipo.	106
Tabla 8. Ejemplo de descriptores secundarios suplementarios generados por transformación así como combinaciones binarias, terciarias y cuaternarias de los descriptores primarios.....	107
Tabla 9. Evaluación de modelos PLS de predicción de contenido de etileno, basados en información de curvas FID/ $^1\text{H-RMN}$ y descriptores matemáticos de decaimiento magnético.....	114
Tabla 10. Evaluación de modelos PLS Predicción módulo de Flexión 1% Secante, basados en información de curvas FID/ $^1\text{H-RMN}$ y descriptores matemáticos de decaimiento magnético.....	125

RESUMEN

El polipropileno (PP) es un derivado en la industria petroquímica, con una producción mundial mayor de 12.5 millones de toneladas por año. La industria de producción de PP invierte gran cantidad de tiempo, entre 24 a 72 horas, para realizar las mediciones de las propiedades físicas y mecánicas a cada lote de PP terminado (ASTM D-4101).

Existe interés por parte de la industria del PP en desarrollar e implementar metodologías analíticas basadas en técnicas espectroscópicas y métodos quimiométricos que proporcionen información fisicoquímica relevante al PP en un menor intervalo de tiempo para una oportuna toma de decisiones con respecto a su calidad y aptitud de uso.

La espectroscopia de resonancia magnética nuclear Protónica (^1H -RMN) es una técnica analíticas no destructiva con capacidad para suministrar información fisicoquímica de interés en la industria de manufactura de resinas plásticas como el polipropileno durante las etapas intermedias del proceso y en el producto final.

La ^1H -RMN permite conocer la composición cristalina, semicristalina y amorfa del polipropileno midiendo las diferentes movilidades en la precesión y relajación de los spines magnéticos moleculares lo cual se relaciona con sus propiedades físico-químicas y mecánicas permitiendo así la predicción de estas propiedades.

Este trabajo presenta el desarrollo de una metodología basada en Resonancia Magnética Nuclear protónica (^1H -RMN) y análisis multivariado, que usando muestras de PP en estado sólido (pellets y polvo) logra determinar el porcentaje de etileno, el módulo de flexión 1% secante y la resistencia a la tensión a la cedencia (Tensile Strength at Yield) en las resinas de PP.

La nueva metodología ^1H -RMN desarrollada proporciona respuestas rápidas, sin preparación de muestra y no requiere el uso de solventes o ningún otro tipo de reactivo nocivos para la salud o el medio ambiente. Por otro lado, mediante el desarrollo de este trabajo se muestra la aplicación del uso de ^1H -RMN en conjunto con las herramientas quimiométricas para uso industrial, por lo cual representa un aporte metodológico nuevo a la comunidad científica nacional e internacional.

1. INTRODUCCIÓN

El polipropileno (PP) es un polímero de la familia de las poliolefinas termoplásticas, que tiene un gran uso como derivado en la industria petroquímica, con una producción mundial mayor de 12.5 millones de toneladas al año, el cual presenta una amplia gama de aplicaciones a nivel industrial y en productos de consumo de la población en general (Bower, 2002). El PP es una poli-olefina que presenta ventajas competitivas frente a muchos polímeros por sus propiedades mecánicas, baja densidad, alta versatilidad y fácil proceso de moldeo así como su relativo bajo costo de producción (Ning, et al.; 2007).

El PP puede ser clasificado en PP homopolímero y PP copolímeros (propileno-etileno) dependiendo de la cantidad de etileno agregado durante el proceso de producción. Los copolímeros de PP están subdivididos en PP copolímero random, copolímero random de bajo etileno y copolímero de impacto. Además de las diferentes clasificaciones de PP existe variación en el peso molecular de la cadena polimérica de PP, la geometría molecular de la cadena de PP y los aditivos agregados para mejorar las propiedades fisicoquímicas del polímero.

La producción industrial del PP requiere una serie de pruebas de control de calidad físicas, mecánicas y químicas para determinar la utilidad del PP en diversos productos. A lote de PP producido le son realizadas pruebas fisicoquímicas para determinar el porcentaje de etileno (%p/p), la tacticidad del PP, el índice de fluidez, resistencia máxima a la tracción, elongación en el punto de cedencia, módulo de flexión, impacto Izod, impacto Gardner entre otras. Las propiedades fisicoquímicas del PP determina el uso y aplicaciones del PP. El nivel de producción de la industria de polimérica está muy relacionado con el tiempo empleado en los análisis químicos y mecánicos para garantizar las características del producto terminado. La industria de producción de PP invierte gran cantidad de tiempo, entre 24 a 72 horas, para realizar las mediciones de las propiedades físicas y mecánicas del PP terminado (ASTM D-4101). Existe interés por parte de la industria del PP en desarrollar e implementar metodologías analíticas basadas en técnicas instrumentales y métodos

quimiométricos que proporcionan información fisicoquímica relevante al PP en un menor intervalo de tiempo.

Espectroscopia de resonancia magnética nuclear Protónica (^1H -RMN) es una técnica analíticas no destructiva que provee información fisicoquímica y es usada en el control de los procesos industriales de producción de polímeros plásticos entre ellos el polipropileno. La ^1H -RMN permite conocer la composición cristalina, semicristalina y amorfa del polipropileno midiendo las diferentes movilidades en la precesión y relajación de los spines magnéticos moleculares (US patent 5,302,897). Los métodos de ^1H -RMN generan un gran volumen de información, relacionada a la estructura molecular de la cadena de PP.

La conversión de datos multivariados en información útil requiere el uso de herramientas matemáticas, estadísticas y herramientas lógicas formales que se han agrupados en los denominados quimiometría disciplina .La quimiometría es una herramienta de análisis multivariado aplicada al estudio de problemas químicos con múltiples variables. Usando quimiometría es posible desarrollar métodos capaces de predecir propiedades físicas-químicas o mecánicas partiendo de técnicas instrumentales multivariadas.

Este trabajo presenta el desarrollo de una metodología basadas en Resonancia Magnética Nuclear protónica (^1H -RMN) usando muestras de PP en estado sólido (pellets y polvo) que permite predecir el porcentaje de etileno, el módulo de flexión 1% secante y la fuerza de tensión a la cedencia (Tensile Strength at Yield) en las resinas de PP. La gran ventaja de este método es que proporciona respuestas rápidas y no requiere el uso de solventes o ningún otro tipo de reactivo que pueden ser nocivos para la salud, por lo tanto esto métodos son considerados ambientalmente limpios o verdes. Por otro lado, mediante el desarrollo de este trabajo se muestra la aplicación del uso de ^1H -RMN en conjunto con las herramientas quimiométricas para uso industrial, por lo cual representa un aporte metodológico nuevo a la comunidad científica nacional e internacional.

2. MARCO TEÓRICO

2.1 POLIPROPILENO

El polipropileno (PP) es un polímero termoplástico de la familia de las poliolefinas, formado de enlaces simples carbono-carbono y carbono-hidrogeno. El PP es el producto de la polimerización donde participa el propileno como monómero que forma las cadenas del polímero (Figura 1). La macromolécula de PP contiene entre 10.000 y 20.000 unidades de monómero (Tripathi, 2002). La densidad promedio del PP es $0,90 \text{ g/cm}^3$ y su estructura molecular consiste de un grupo metilo ($-\text{CH}_3$) unido a una cadena lineal polimérica. Por medio del arreglo molecular del grupo metilo se logran obtener diferentes configuraciones estereoquímicas en la polimerización del propileno (Tripathi, 2002); (Carraher, 2008).

El PP pertenece a la familia de los termoplásticos semicristalinos, su gran desarrollo comercial se dio a partir de 1950 luego del descubrimiento de los catalizadores de Ziegler-Natta (Harding y van Reenen, 2011). Los catalizadores Ziegler-Natta están basados principalmente en sales de metales de transición como el TiCl_3 y TiCl_4 , junto con cocatalizadores de alquilaluminio como el trietilaluminio ($\text{Al}(\text{CH}_2\text{CH}_3)_3$) que permiten la producción de esta resina en condiciones no extremas de temperatura y presión. Los pasos básicos del mecanismo de polimerización del PP son presentados en la figura 2.

Simultáneamente durante la síntesis del PP es posible el control morfológico y de estereoregularidad del material y por consiguiente las propiedades del mismo son modificadas de acuerdo a su uso. La estereoselectividad de los catalizadores de Ziegler-Natta es clave en determinar ciertas propiedades del PP, puesto que la posición del grupo metilo lateral produce centros quirales de acuerdo al plano de simetría dado por la cadena principal (Harding y van Reenen, 2011). Esto produce tres configuraciones espaciales que resultan en diferentes propiedades físicas y mecánicas del PP. Las diferentes

configuraciones estereoquímicas del PP permiten clasificarlo como isotáctico, sindi táctico y atáctico (Tripathi, 2002).

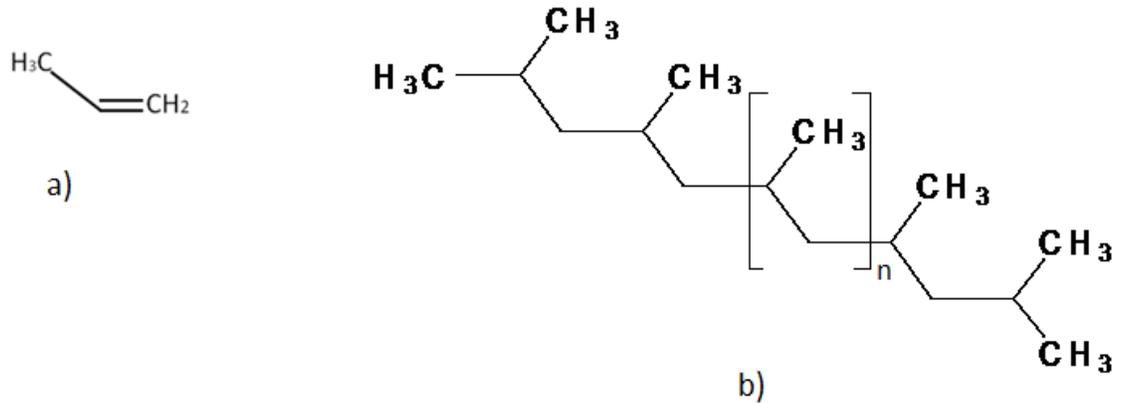


Figura 1. Formulas estructurales a) representación del monómero de propileno, b) representación del polipropileno.

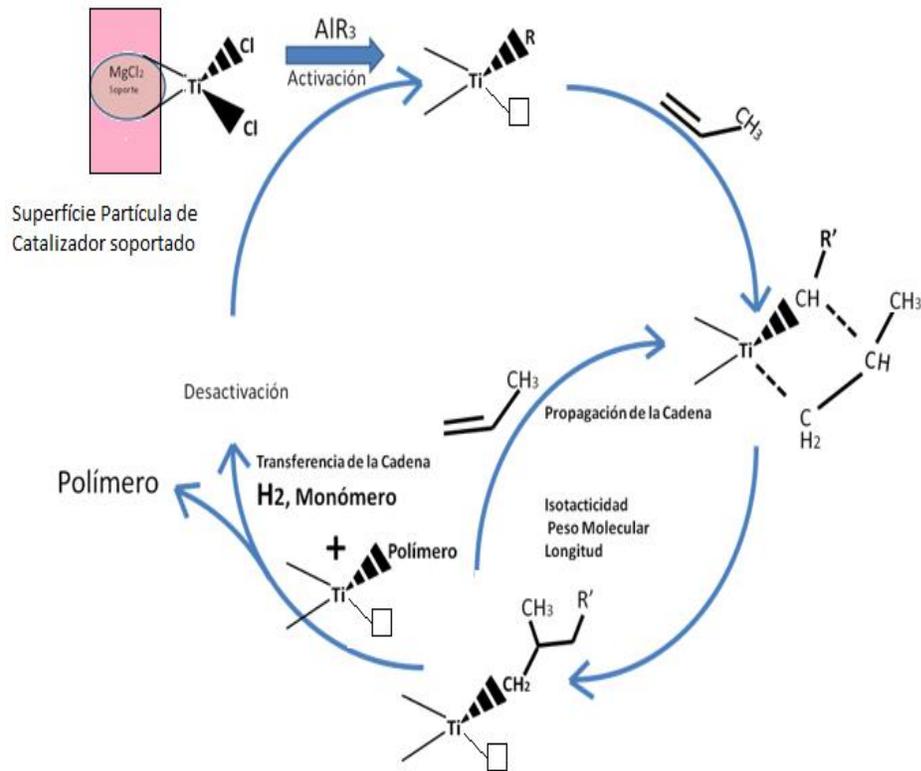


Figura 2. Esquema que muestra el ciclo de polimerización del polipropileno.

- **Configuración del polipropileno isotáctico, PP-i:** En esta configuración la disposición de todos los grupos metilos laterales es por encima o por debajo del plano de la cadena principal.
- **Configuración del polipropileno atáctico, PP-a:** Los grupos metilos laterales no guardan una disposición regular por encima o por debajo del plano de la cadena principal, sino aleatoriamente.
- **Configuración del polipropileno sindiotáctico, PP-s:** Configuración en donde todos los grupos metilos laterales están dispuestos de manera alternada por encima y por debajo del plano de la cadena principal (O dian, 2004). Estas estructuras se presentan en la Figura 3.

Las moléculas de PP con configuración isotáctica y sindiotáctica son las únicas capaces de formar estructuras cristalinas (Bower, 2002). Las propiedades mecánicas del PP dependen en gran medida de la proporción entre regiones cristalinas y amorfas en la estructura del termoplástico sólido. La regularidad del PP Isotáctico determina que este pueda cristalizar, lo cual impacta las propiedades físicas de este polímero. Las configuraciones isotácticas y sindiotácticas, debido a su gran regularidad, tienden a adquirir en estado sólido una disposición espacial ordenada, semicristalina, que confiere al material unas propiedades físicas excepcionales (Carragher, 2008). La forma atáctica, en cambio, no tiene ningún tipo de cristalinidad, es totalmente amorfa, por lo cual es menos usada en la industria que las anteriores. Comercialmente las plantas de producción de PP están diseñadas para producir PP homopolímero isotáctico con un bajo nivel de la configuración atáctica (Ning, et al.; 2007). El PP sindiotáctico no se produce comercialmente en masa (Tripathi, 2002).

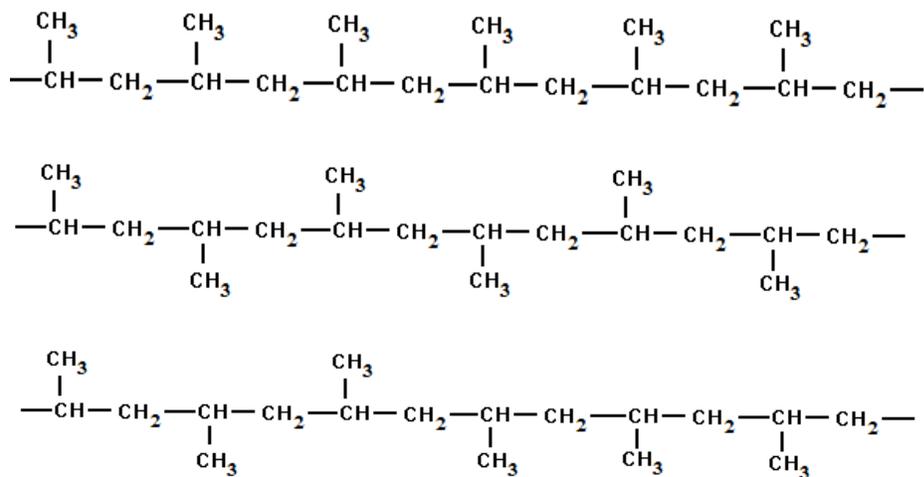


Figura 3. Configuración estereoquímica del PP a) configuración isotáctica b) configuración sindiotáctica c) configuración atáctica.

Las plantas de producción de PP son capaces de producir PP con diferentes niveles de estereoregularidad y por consiguiente diferentes niveles de cristalinidad. A nivel de laboratorio se realiza una prueba que permite determinar el valor aproximado del contenido de cadenas atácticas en una muestra de PP. La prueba es denominada solubles en xileno a temperatura ambiente. El PP de configuración atáctica es soluble en xileno a temperatura ambiente, las configuraciones sindiotáctica y atáctica del PP no son solubles en xileno a temperatura ambiente. Las tres configuraciones de PP, PP atáctico, PP sindiotáctico y PP isotáctico son solubles en xileno a temperatura de 140°C (Harding y van Reenen, 2011). Esta es una prueba normalizada y una descripción detallada de la misma se encuentra en la norma ASTM D-5492 Standard Test Method for Determination of Xylene Solubles in Propylene Plastics (ASTM D-5492 - 1998). Comercialmente se producen PP con valores de solubles en xileno desde cerca de 1% hasta aproximadamente 20%. El control preciso de estos rangos se logra con la selección de catalizadores especialmente desarrollados y con ajustes de condiciones de polimerización adecuadas.

Los métodos estandarizados que se aplican en la industria de manufactura de resinas plásticas desde la obtención de los especímenes de ensayos hasta la realización de los mismos son entre otros los de ASTM, American Society for Testing and Materials los

cuales tienen sus equivalentes métodos ISO, International Standard Organization. La tabla 1 presenta un listado de los métodos ASTM empleados en la producción del PP.

Tabla 1. Normas ASTM para determinar propiedades físico-químicas y mecánicas en el polipropileno.

Norma	Descripción
ASTM D-792	Test Methods for Density and Specific Gravity (Relative Density) of Plastics by Displacement
ASTM D-1238	Test Methods for Flow Rates of Thermoplastics by Extrusion Plastometer
ASTM D-5492	Test Methods for Determination of Xylene Solubles
ASTM D-638	Test Method for Tensile Properties of Plastics
ASTM D-790	Test Methods for Flexural Properties of Unreinforced and Reinforced Plastics and Electrical Insulating Materials
ASTM D-256	Test Methods for Determining the Izod Pendulum Impact Resistance of Plastics
ASTM D-785	Test Method for Rockwell Hardness of Plastics and Electrical Insulating Materials
ASTM D-1525	Test Method for Vicat Softening Temperature of Plastics
ASTM D-648	Test Method for Deflection Temperature of Plastics Under Flexural Load in the Edgewise Position
ASTM D-6110	Test Method for Determining the Charpy Impact Resistance of Notched Specimens of Plastics
ASTM D-5420	Test Method for Impact Resistance of Flat, Rigid Plastic Specimen by Means of a Striker Impacted by a Falling Weight (Gardner Impact)
ASTM D-3835	Test Method for Determination of Properties of Polymeric Materials by Means of a Capillary Rheometer

2.2 COPOLIMERIZACIÓN DE MOLÉCULAS DE POLIPROPILENO

Podemos encontrar el PP en formas pura (homopolímero) y modificados (copolímeros y/o con aditivos agregados en concentraciones pequeñas o apreciables). Los copolímeros del

PP son clasificados de acuerdo a la materia prima o tipos de monómeros usados en su fabricación, como se muestra a continuación.

Polipropileno homopolímero. El PP homopolímero (PP-h) está formado de cadenas poliméricas de material puro sin agregados monoméricos distintos al propileno (Figura 1) o aditivos que puedan cambiar sus propiedades fisicoquímicas. Lo anterior significa que el monómero de PP mencionado es la unidad repetitiva a lo largo de las cadenas de este polímero. El PP-h es una resina altamente cristalina, con alta temperatura de fusión (164°C), con una Temperatura de Transición Vítrea (Tg) de aproximadamente. 0°C. El PP-h posee buenas propiedades dieléctricas, su resistencia a la tensión es excelente en combinación con la elongación (permitiendo ser orientado de manera relativamente sencilla), presenta apariencia translúcida, excelente resistencia a altas temperaturas y buena resistencia a diversos productos químicos (Carraher, 2008). Las moléculas de PP-h son empleados ampliamente para extrusión de lámina, envases soplados, tubería, fibra, multifilamentos e inyección de alta rigidez, entre otros. No es recomendado para aplicaciones cuyo uso es a temperaturas inferiores a 0°C, en cambio es recomendado para aplicaciones de llenado en caliente. Sus usos más comunes son inyección de artículos de uso doméstico, envases y contenedores, muebles, juguetes y tapas, extrusión de fibras y filamentos para cordeles, alfombras, hilos, tapices, sacos, pañales desechables, extrusión de grandes planchas (hasta 250 μm de espesor) para cubiertas de muebles y estanques y planchas delgadas para termoformar envases y contenedores.

Adicionalmente a las cadenas de PP-h, existen dos familias de copolímeros o terpolímeros que son formadas por los monómeros de propileno y etileno y/o butano en un amplio rango de concentraciones. Estas resinas sintéticas son denominadas copolímeros de polipropileno/etileno random (CPE-r), copolímeros de impacto polipropileno/etileno (CPE-i) y copolímeros de polipropileno/etileno de bajo contenido de etileno (CPE-LER).

Los CPE-r (Figura 4) típicamente contienen hasta 6 %p/p de etileno o butano que hace parte de la cadena polimérica insertándose en forma aleatoria a lo largo del polímero

(Carraher, 2008). La presencia del monómero de etileno en la cadena polimérica permite la formación de defectos que reducen la estéreo-regularidad y por consiguiente la tacticidad, lo cual se traduce en disminución de la cristalinidad y desplazamiento del punto de fusión hacia valores por debajo del punto de fusión del PP-h (Carraher, 2008).



Figura 4. Copolímero random de PP. P = Propileno, E = Etileno.

Dependiendo del contenido de copolímero en la cadena de PP ocurre una modificación de las propiedades mecánicas del copolímero (Bower, 2002). Como consecuencia en la disminución de la cristalinidad también se consigue un aumento de la claridad o disminución de la opacidad (Carraher, 2008). Los cristales formados en el CPE-r son menos perfectos y su densidad es menor comparada con la densidad de los cristales de PP-h (Tripathi, 2002). Estos cristales poseen un índice de refracción más bajo y cercano al índice de refracción de la fase amorfa por lo que la luz sufre una menor refracción (Tripathi, 2002). Este factor hace que los objetos vistos a través de una película o lámina hecha con CPE-r se vean con mayor claridad y menor opacidad con bordes mejor definidos. Los CPE-r son usados principalmente en aplicaciones para empaque y por su menor punto de fusión en aplicaciones donde se necesita un bajo umbral de selle en la misma industria de empaques. La disminución de la cristalinidad a su vez disminuye la propiedad mecánica conocida como el módulo de flexión que es una medida de la rigidez de los materiales. La tabla 2 muestra datos típicos de las propiedades mecánicas que diferencian a los tres tipos de familias de polipropileno.

El otro tipo de PP modificado es el CPE-i, también conocido como copolímeros heterofásicos. Los CPE-i consisten de una fase o bloque copolímero propileno-etileno en proporción cerca de 1:1 (fase de caucho) dispersada sobre una fase de PP-h que constituye el núcleo. Los niveles o contenidos de la fase de caucho están en el rango de 6-38%.

Tabla 2. Propiedades mecánicas típicas de las diferentes familias de polipropileno

Familia de Polipropileno	Rata de flujo del fundido (MFR) a 230°C [g / 10 min] ¹	Módulo de Flexión a 1,3 mm/min [psi] ²	Resistencia al impacto Izod ranurado a 23°C [pie-Lb/pulg] ³	Dureza Rockwell, [Escala R] ⁴	Temperatura de deflexión térmica bajo carga a 66 psi [°C] ⁵
Homopolímero	20	348091	-----	----	----
	20	290075	0,60	104	124
	40	274847	0,60	100	121
	12	274847	0,60	105	118
	2	259618	2,62	100	110
	4	249465	0,80	99	97
	20	249465	0,60	104	100
	12	240037	0,50	97	99
	100	229885	-----	----	----
	0,5	229885	2,99	95	93
	70	219732	-----	----	90
	1,5	214656	1,31	----	87
	35	189999	0,60	98	90
	Copolímero random	3	219732	0,69	95
25		164618	-----	94	98
35		164618	0,80	80	87
10		164618	1,01	88	91
1		139961	1,50	80	90
25		129809	0,69	80	85
10		124732	0,69	80	85
6		119656	0,90	80	85
2,5		100076	1,57	65	70
7		84847	1,50	65	65
Copolímero de impacto	6	71794	-----	----	----
	35	189999	1,31	85	90
	4	189999	2,52	82	81
	0,5	179847	10,10	75	90
	50	169694	1,03	90	117
	4	150114	10,10	65	75
	11	145038	2,52	72	77
	2	139961	12,16	61	72
	20	139961	3,46	67	77
	30	139961	1,03	80	80
	6	134885	12,16	70	90
	16	124732	12,16	40	80
	1	124732	12,16	55	79
20	114580	10,10	65	85	

Fuente.: (Edward P. Moore, 1996).

1. ASTM D 1238; 2. ASTM D 790; 3. ASTM D 256; 4. ASTM D 785; 5. ASTM D 648

La aplicación principal de los CPE-i es la fabricación de artículos con alta resistencia al impacto a baja temperatura (Tripathi, 2002). La distribución del etileno en la fase dispersa y en el polímero total es descrita mediante los siguientes términos:

- 1) **Etileno total (Et):** Se refiere al contenido total de etileno en peso en la cadena polimérica, el cual se encuentra entre 3-19% w/w.
- 2) **Etileno en el copolímero (Ec):** Hace referencia al contenido de etileno en copolímero de propileno-etileno o fase caucho, el cual se encuentra entre 40-60 % w/w.
- 3) **Fracción del copolímero (Fc):** Es la porcentaje de fase copolímero o caucho en el polímero total, cuyos valores oscilan entre 6-38 % w/w.

La composición, peso molecular, cantidad y morfología de las fases dispersada (caucho propileno-etileno) y la fase continua (homopolímero) es crucial para tener un copolímero de impacto efectivo y con un deseable balance de propiedades de resistencia al impacto y buena rigidez simultáneamente (Tripathi, 2002).

La figura 5 muestra el efecto de la variación en el peso molecular representado por el MFR sobre la resistencia mecánica al impacto.

La capacidad de cristalizar la cadena es un factor crucial que rige la morfología final del polímero y a su vez la cristalinidad depende principalmente de la tacticidad de las cadenas poliméricas.

Durante la producción industrial el grado de tacticidad se puede variar para conseguir diferentes grados de cristalinidad que pueden ser hasta el 70% en el caso del PP homopolímero isotáctico. Para alcanzar esos altos niveles de cristalinidad se requieren cadenas largas con secuencias isotácticas que formen cristales, la resonancia magnética nuclear protónica ($^1\text{H-RMN}$) es suficientemente sensible para detectar y dar una señal

diferente para los distintos grados de cristalinidad debido a que la movilidad de los protones a nivel molecular varía dependiendo del grado de cristalinidad.

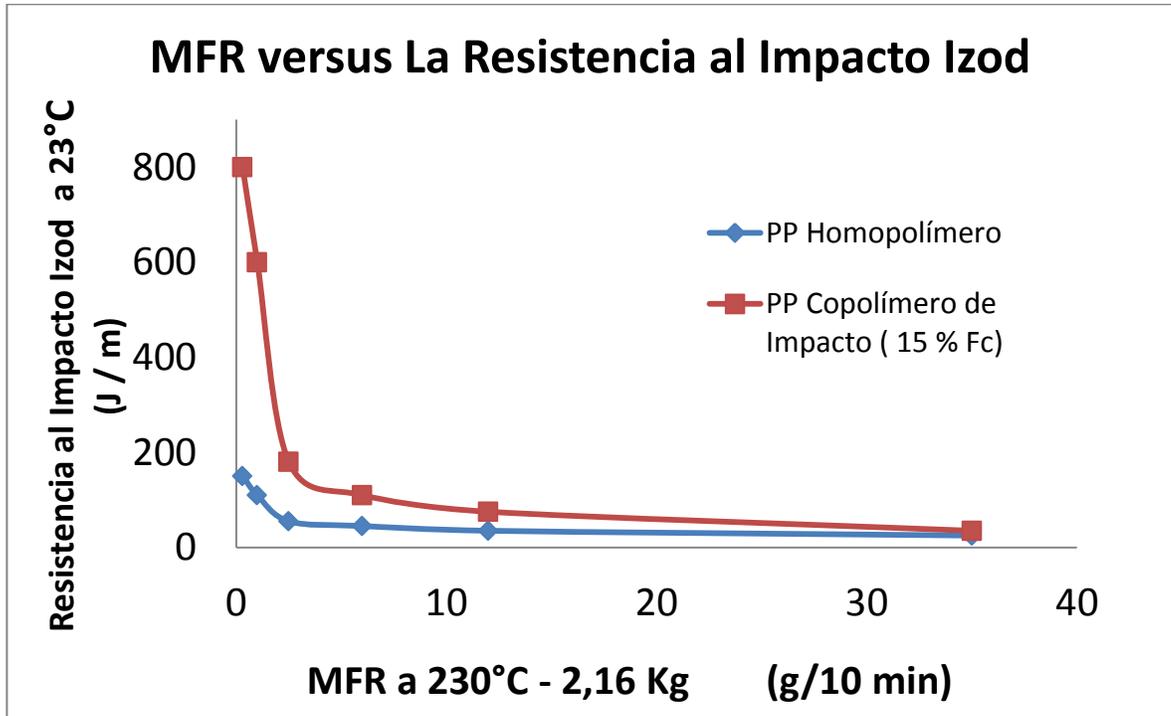


Figura 5. Peso molecular, representado indirectamente por la rata de flujo de fundido - MFR (ASTM D 1238), sobre la resistencia al impacto Izod (ASTM D 256). Fuente:(Edward P. Moore, 1996).

La cristalinidad determina el nivel de propiedades mecánicas del PP, por lo tanto la técnica de $^1\text{H-RMN}$ puede suministrar información de la resistencia mecánica del polímero al censar la cristalinidad o grado de movilidad de los protones unidos a cadena isotácticas y a cadenas atácticas (Edward P. Moore, 1996).

Durante el procesamiento de PP también puede ser variado su grado de cristalinidad al variar algunos parámetros. Entre los parámetros de procesamiento importantes tenemos, las condiciones de fundido, la orientación, historia térmica, la adición de aditivos, agentes nucleantes y clarificantes que ejercen su efecto en pequeñas cantidades, en el orden de partes por millón (Arroyo, et al.; 1997) así como reacciones de post-síntesis relacionadas

con degradación del peso molecular del polímero y el entre cruzamiento pudiendo ser detectados todos estos efectos y variaciones por la técnica ^1H -RMN de baja resolución mediante la interpretación adecuada de la curva de relajación (FID; Free Induction Decay) (Fink, et al.; 2000). El entendimiento de los efectos de estas variables estructurales y de procesamiento del polímero sobre la morfología del PP es esencial, porque es la morfología la que suministra información directa sobre las propiedades del enlace del polímero siendo aún más fácil acceder a esta información de la resistencia mecánica del PP de forma rápida y sencilla mediante el ^1H -RMN para el control de calidad de la producción y liberación de los lotes finales producidos como alternativa a las metodologías tradicionales de la ASTM que resultan muy tediosas y laboriosas consumiendo mucho tiempo (Tripathi, 2002).

La opacidad o de manera contraria se puede denominar como el grado de transparencia de una pieza fabricada con polipropileno depende entre otros factores del grado de la naturaleza o familia del polipropileno así como de la cristalinidad y aditivos nucleantes y clarificantes; el mayor nivel de cristalinidad presente en el PP-h explica su mayor opacidad comparado con un PP-r, la Tabla 3 ilustra el efecto de la cristalinidad sobre las propiedades ópticas del polipropileno.

El proceso de cristalización natural del polipropileno isotáctico, PP-i, es variado industrialmente por la adición de aditivos específicos llamados agentes nucleantes los cuales favorecen una cristalización heterogénea, los agentes nucleantes acortan el tiempo de inducción de cristalización del polímero, suministrando superficies foráneas o núcleos que reducen la energía libre de formación de los núcleos del polímero, estas regiones cristalinas aumentadas en el PP-i darían una curva característica de FID del RMN de baja resolución. algunos ejemplos agentes nucleantes benzoato de sodio, talco, caolín y sales del ácido fosfórico (P. Zhang, Liu, & Li, 2006).

Tabla 3. Efecto de la naturaleza del PP y de la presencia de agente clarificante sobre la propiedad óptica opacidad.

Familia de Polipropileno	Opacidad ¹ (Haze)%	
	Sin aditivo clarificante	Con aditivo clarificante
Homopolímero	60 – 65	10 – 15
Copolímero random	40 – 45	8 – 12

Fuente: Milliken & Co. (Edward P. Moore, 1996).

1. ASTM D 1003.

2.3 PROPIEDADES MECÁNICAS DEL POLIPROPILENO

Para la determinación de las propiedades mecánicas en el polipropileno por los métodos tradicionales el primer paso consiste en transformar los pellets en especímenes de forma estándar para el ensayo mediante el proceso de moldeo por inyección o extrusión (Bower, 2002).

2.4 PROCESOS DE TRANSFORMACIÓN

Los procesos de transformación en la terminología de polímeros, se les denominan en general, a aquellas operaciones cuya finalidad consiste en fabricar objetos o artículos elaborados o semi-elaborados tales como láminas o preformas a partir de los polímeros base, que están generalmente en forma de polvo, pellets o gránulos (Carraher, 2008).

Los procesos de transformación se pueden dividir en dos grandes grupos en función del principio de las operaciones de moldeo: continuos y cíclicos. En los procesos continuos, se crea un flujo constante de material fluido, al que se da forma al ser extruido a través de un

orificio o boquilla (extrusión) o al hacerlo pasar entre rodillos (calandrado). En los procesos cíclicos el material fundido rellena un molde o cavidad por la acción de una presión mecánica, como en el caso del moldeo por compresión, o hidrostática como en el moldeo por inyección.

En la etapa de previa al moldeo en la mayoría de procesos de transformación, ya sean sencillos o combinados, suele intervenir una o más extrusoras de tornillo que permiten calentar y homogeneizar el polímero con sus aditivos de manera eficiente. La elección de un determinado proceso de transformación para un artículo vendrá condicionada por diversas variables, como son el diseño, el volumen de producción, la calidad y el precio final del producto. En función del proceso de transformación empleado, se elegirán como materias primas polímeros de diferentes grados, que se distinguen principalmente por su fluidez escogiendo los más viscosos para procesos de extrusión y conformado y los más fluidos para los de inyección.

Tabla 4. Algunos procesos de transformación y aplicaciones de uso final del Polipropileno.

Aplicaciones de Uso Final	Procesos de transformación del PP	Industrias en Colombia (Ciudad)	Homopolímero	Copolímero random	Copolímero de Impacto
Película (Film)	Película de PP Bi-Axialmente Orientada (BOPP Film, Bi-axially Oriented Polypropylene Film)	Biofilm (Cartagena)	X	X	
	Película fundida convertida en el proceso de rodillo de enfriamiento. (Cast Film, Film converted in the chill roll process)	Películas Extruidas, S.A. (Bogotá)	X	X	X
	Película Tubular Enfriada por Baño de Agua (TWQ, Tubular Water Quenched Film)	Polybol S.A (Cartagena)	X	X	
Textiles	Monofilamento y Rafia	Ciplas S.A (Bogotá)	X	X	
	Textiles de Filamento Continuo & Filamento Continuo Abultado (Staple CF & BCF, Continuous Filament & Bulked Continuous Filament)	Bonlan Andina (Cali)	X		
	Tela no tejida de polipropileno (Spun Bond)	Bonlan Andina (Cali)	X	X	
	Fundido soplado (Melt Blown)	Bonlan Andina (Cali)	X		
Productos Industriales y Durables	Moldeo Por Inyección	Corplas S.A. (Cartagena)	X	X	X
	Extrusión	Ajover; Polyban (Cartagena)	X	X	X
Empaques Rígidos	Termoformado	Multidimensionales S.A. (Bogotá)	X	X	X
	Moldeo por soplado	Corplas S.A. (Cartagena)	X	X	X

Finalmente, cabe reseñar que el proceso de transformación también afectará a la historia termo-mecánica que sufre el plástico, afectando su estructura y morfología, y que luego se verá reflejada en el cambio de las propiedades físicas y mecánicas finales del producto.

Una recopilación de las aplicaciones modernas del polipropileno y los procesos de transformación usados en la industria colombiana para obtener los diferentes productos finales son resumidos en la tabla 4.

2.5 PROPIEDADES MECÁNICAS

Debido a que en casi todas las condiciones de uso final de la mayoría de las aplicaciones del polipropileno se ejerce algún grado de carga o presión sobre el artículo fabricado, las propiedades mecánicas se constituyen frecuentemente en las más importantes y que determinan el éxito de una aplicación. Por lo tanto la selección de un material plástico con las propiedades mecánicas adecuadas determina en gran medida el éxito del artículo fabricado bajo los requerimientos de uso final. Entre las diferentes propiedades mecánicas realizadas en los laboratorios bajo condiciones estandarizadas de ensayo tenemos, la resistencia a la tensión, la elongación, el módulo de flexión, y la resistencia al impacto.

La determinación de las propiedades mecánicas se base en la obtención de la curva de esfuerzo-deformación (stress-strain) al aplicar a un espécimen de ensayo normalizado una fuerza en modo tensión, flexión, compresión o torsión (Figura 6).

El tramo inicial de la curva esfuerzo-deformación entre los puntos A y C, es recto, y mantiene una relación lineal entre el esfuerzo aplicado y la deformación resultante, en este tramo, la deformación mecánica del polipropileno cumple la ley de Hooke, la cual establece que para un material elástico el esfuerzo es proporcional a la deformación. En punto C, es donde la gráfica empieza a desviarse de la línea recta, este punto es llamado límite de proporcionalidad. Por debajo de este punto el comportamiento del polipropileno es elástico siendo la deformación recuperable es decir que al cesar el esfuerzo, la pieza regresa a su dimensión original.

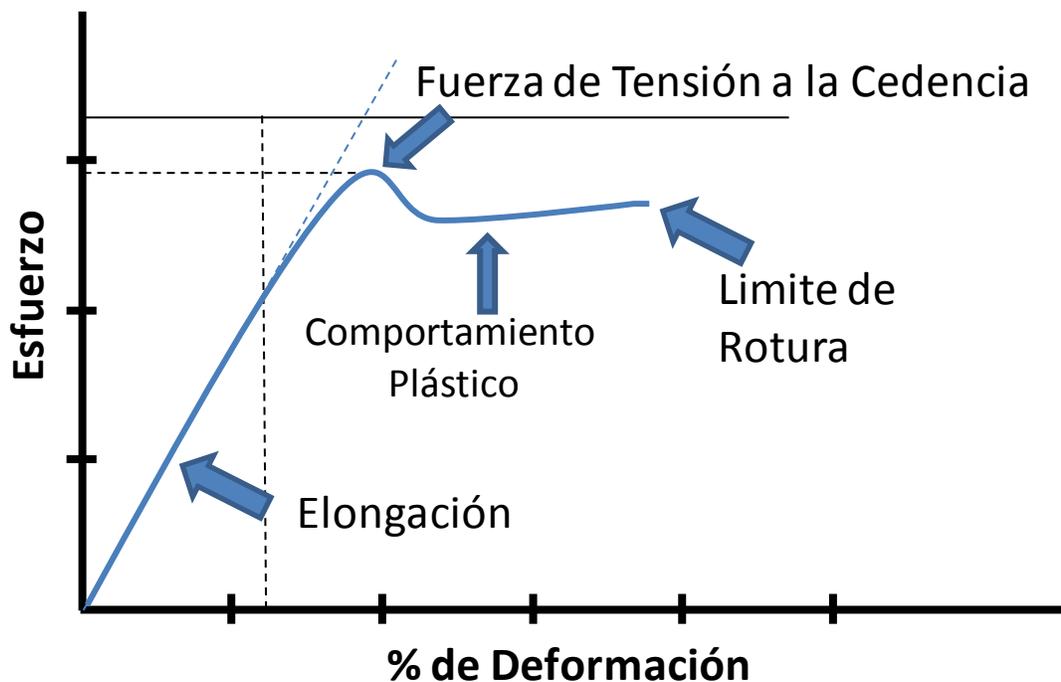


Figura 6. Gráfica esfuerzo-deformación característica de un espécimen normalizado de Polipropileno, generada en un ensayo para determinar las propiedades mecánicas. (Shah, 1984).

En la literatura general se explica que la deformación hasta el punto B en la figura 6 es relativamente pequeña y ha sido asociada con la flexión y estiramiento de los enlaces carbono-carbono a lo largo de la cadena hidrocarbonada de polipropileno. Este tipo de deformación es instantánea y recuperable, no hay desplazamiento permanente de las moléculas de polipropileno con respecto a las otras (Figura 6). La deformación que ocurre más allá del punto C en la Figura 6, es similar al enderezamiento o estiramiento de las cadenas moleculares de polipropileno enrolladas o embobinadas sobre sí mismas (Figura 6). No hay deslizamiento intermolecular en las cadenas poliméricas, es importante destacar que la deformación puede ser recuperable al final pero no instantáneamente. Si el polipropileno es sometido a deformaciones más allá del Yield Point o del límite elástico del polímero no son recuperables (Figura 6). Esta deformación se da porque ocurre un desplazamiento real de las cadenas de polipropileno con respecto a las otras. Las moléculas poliméricas desplazadas no pueden deslizarse y regresar a su posición original y por lo

tanto ocurre una deformación permanente. Estos tres tipos de deformaciones descritas en la Figura 6, no ocurren separadamente sino que se dan superpuestas unas con otras. La flexión y estiramiento de los enlaces carbono-carbono a lo largo de la cadena de polipropileno, ocurre de forma casi instantánea. Sin embargo el proceso de des enrollamiento de las cadenas de polipropileno es relativamente lento. El deslizamiento molecular de las cadenas poliméricas es el efecto más lento de los tres procesos de deformación.

En la literatura también se explica la deformación del polipropileno a partir de un modelo mecánico que duplica el comportamiento de los plásticos bajo varias condiciones externas. Modelo mecánico usando resortes y amortiguadores conocido como el modelo de Maxwell(Vishu, 1984) (Figura 7).

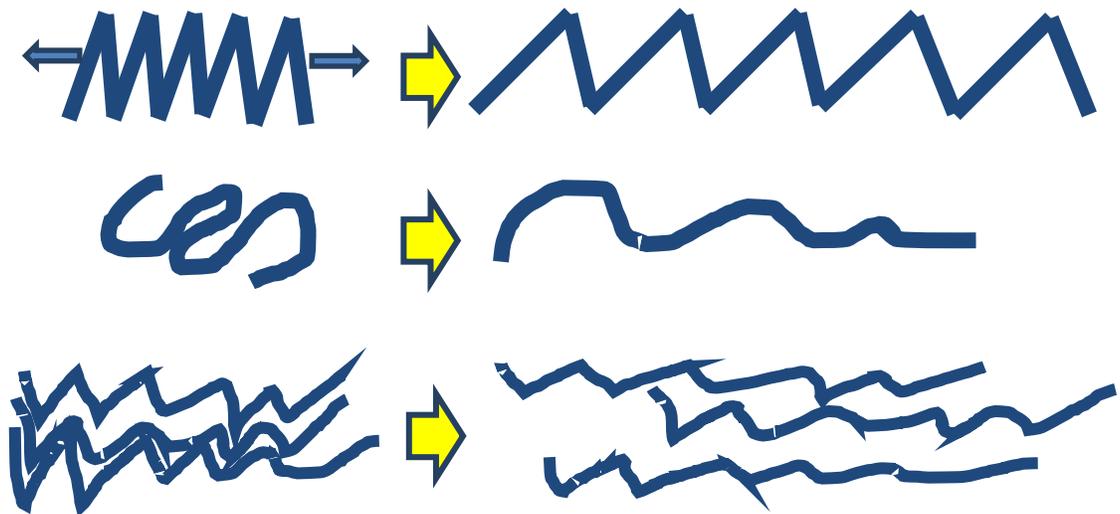


Figura 7. Diferentes tipos de extensión a nivel molecular del polipropileno durante la deformación mecánica. Tomado con modificación de Shah,1984.

Los especímenes de ensayo normalizados para realizar las pruebas de resistencia a la tensión y módulo de flexión son preparados por el proceso de moldeo por inyección a

partir de las muestras en forma de pellets usando las condiciones y guías establecidas en la Norma ASTM D-4101. Las muestras obtenidas son especímenes tipo I en forma de campana (dumb-bell-shaped specimens) (Figura 8), con las características y dimensiones descritas en la norma ASTM D-638 el cual es el espécimen comúnmente usado para el ensayo de materiales plásticos rígidos y semi-rígidos.

Debido a que las propiedades de tensión para algunos plásticos y en especial para el polipropileno varían con pequeños cambios en la temperatura, los ensayos fueron realizados en condiciones estándar de laboratorio. Los especímenes una vez obtenidos fueron acondicionados por al menos 40 horas a $23 \pm 2^\circ\text{C}$ y $50 \pm 5\%$ de humedad relativa. Para el análisis de las propiedades mecánicas son empleadas máquinas universales de medidas tipo Instron en diferentes versiones (ASTM D-4101).

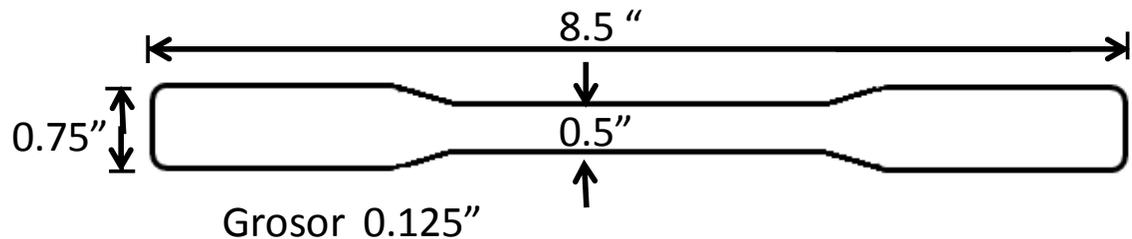


Figura 8. Espécimen de ensayo para la prueba de tensión y flexión.

2.5.1 Resistencia a la tensión a la cedencia (Tensile strength at yield)

Las propiedades de tensiones son unas de las más importantes y ampliamente especificadas propiedades de los materiales plásticos puesto que son una indicación de la fortaleza del material. La prueba de tensión, en forma general es la medida de la habilidad de un material de soportar fuerzas que tienden a estirarlo y determina hasta que punto se estira antes de romperse. El módulo de tensión del polipropileno es una indicación de la rigidez (stiffness) relativa del material y puede ser determinado a partir de la curva de esfuerzo-deformación tal como se explicó anteriormente.

La resistencia a la tensión (Tensile) es determinada siguiendo la norma ASTM D-638 usando una máquina universal de medidas Instron Modelo 3345 con una distancia entre mordazas de 101.6 mm y una velocidad de cruceta de 50 mm/min. Antes de realizar el ensayo a cada espécimen se le mide el espesor y se analizan cinco especímenes por muestra. La resistencia a la tensión a la cedencia fue calculado de la siguiente forma:

$$\text{Resistencia a la tensión a la cedencia psi} = \frac{\text{Máxima carga registrada (Lb)}}{\text{Área de la sección transversal (Pulg)}^2} \text{ (EC. 1)}$$

2.5.2 El módulo de flexión 1% secante

El conocimiento del comportamiento en modo flexión del polipropileno es de mucho interés y es una propiedad muy importante cuyo valor determina en gran medida el valor del producto porque determina la aplicabilidad de la resina para los diseñadores de piezas y para los convertidores.

La resistencia a la flexión es la habilidad de un material para soportar las fuerzas de flexión aplicadas perpendicularmente a su eje longitudinal. Las tensiones inducidas debido a la carga de flexión, son una combinación de tensiones de compresión y de tracción, la Figura 9 ilustra este efecto.

Fuerzas envueltas en la flexión de un espécimen

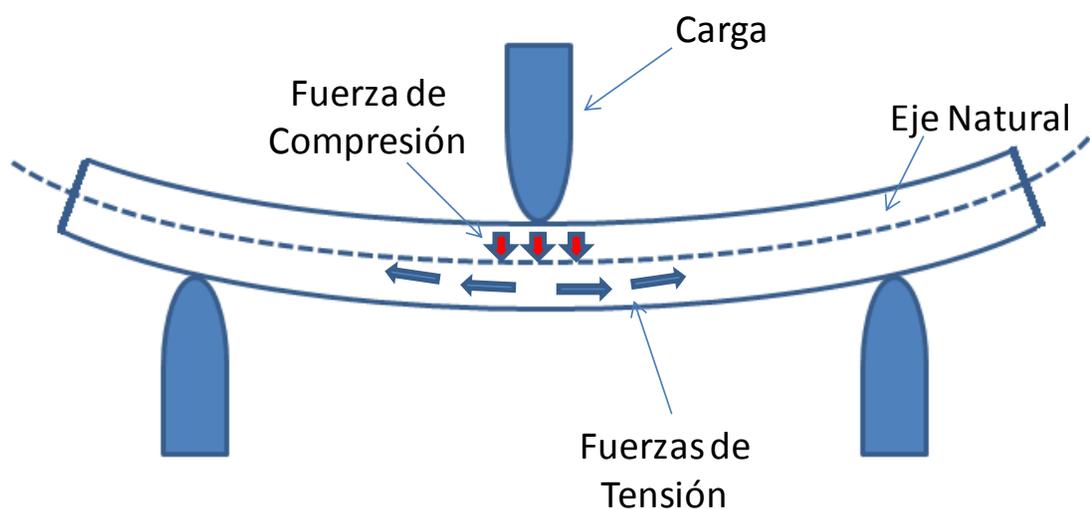


Figura 9. Esquema de las fuerzas actuando sobre la flexión de un espécimen de polipropileno. (Tomado con modificaciones de McGraw-Hill Book Company).

Es determinado según lo establecido en la norma ASTM D790 procedimiento A, usando la maquina universal de medida Instron 3345. Para esto son empleadas tres especímenes por cada muestra y los resultados son reportados como promedio de las tres medidas usando una velocidad de aplicación de la carga de aproximadamente 1.3 mm/min.

Las propiedades mecánicas en flexión fueron calculadas y reportadas en términos de los máximos esfuerzos y deformaciones que ocurren en la superficie exterior del espécimen de ensayo. Una explicación detallada de los cálculos realizados para obtener el valor del módulo de flexión 1% secante se encuentra en la norma ASTM D 790.

2.6 ANÁLISIS EN MUESTRAS SOLIDAS DE POLIPROPILENO USANDO CURVAS DE DECAIMIENTO MAGNÉTICO INDUCIDO (FID) POR RESONANCIA MAGNÉTICA NUCLEAR PROTÓNICA (^1H -RMN) DE BAJA RESOLUCIÓN.

La espectroscopia de resonancia magnética nuclear Protónica (^1H -RMN) es una técnica que se usa cada día con más frecuencia para cuantificar en química analítica, desarrollado en 1946 por dos grupos de investigadores en los EE.UU., de la Universidad de Stamford y del Instituto Tecnológico de Massachusetts (MIT). La tecnología de radar desarrollado durante la Segunda Guerra Mundial hizo posible muchos de los aspectos electrónicos de la ^1H -RMN. Con el desarrollo de nuevos hardware de análisis de información, los químicos empezaron a aplicar la tecnología a los problemas de la química y la física (Duer, 2002).

La ^1H -RMN es usada en el control de los procesos industriales de producción de polímeros plásticos entre ellos el polipropileno. La ^1H -RMN permite conocer la composición cristalina, semicristalina y amorfa del polipropileno midiendo las diferentes movilidades en la precesión y relajación de los spines magnéticos moleculares, luego de aplicar la resonancia magnética nuclear (US patent 5,302,897).

En el análisis de PP homopolímero y copolímeros, el ^1H -RMN opera bajo el principio que cada partícula tiene un spin magnético resultante. Si el número de spines en un núcleo es par los espines se cancelan. Si el número de spines nucleares es impar, se produce un spin resultante diferente de cero, en el caso de los polímeros de PP núcleos de los protones tienen un spin resultante diferente de cero (Choi y Kerr, 2003).

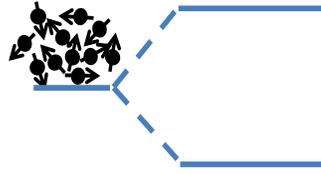
En los átomos y moléculas de muestras poliméricas, los momentos magnéticos generados por los spines atómicos de los átomos de hidrógeno están cuantizados y orientados aleatoriamente. Cuando un campo magnético externo es aplicado a las muestras del polímero, los spines atómicos de los núcleos de hidrógeno se orientan en la dirección del campo magnético externo. Los spines de menor energía se alinean en dirección del campo magnético aplicado, los espines de mayor energía se alinean en dirección opuesta al campo magnético aplicado. Figura 10.

Dependiendo de la intensidad del campo magnético externo aplicado a las partículas de polímeros, es producida una distribución de espines de alta y baja energía (Macomber,

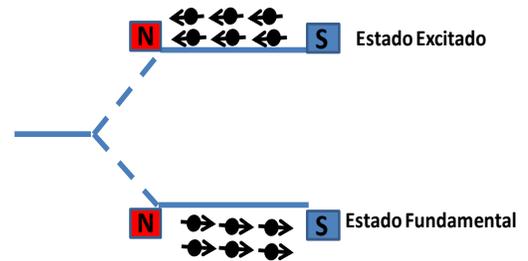
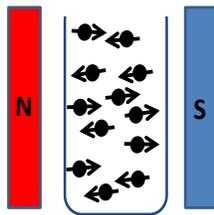
1998). La diferencia de energía entre los dos tipos de espines ($1/2$ y $-1/2$) depende de la intensidad del campo magnético aplicado y el entorno químico de cada átomo en la partícula polimérica. La frecuencia de precesión para los espines magnéticos depende de la intensidad del campo magnético aplicado (Keeler, 2002). Un pulso de radio frecuencia (RFP) con la frecuencia de precesión de los espines de interés, tienen la energía necesaria para invertir la distribución de las poblaciones de los espines de los núcleos atómicos de los núcleos de hidrógeno en el polímero. Cuando es aplicado el proceso de excitación por radio frecuencia, un número de espines de baja energía absorben y el vector resultante de la suma de los espines magnéticos se invierte mientras dura la excitación de la radiofrecuencia (Duer, 2002). Al retirar el RFP se produce una emisión de radiofrecuencia por parte del spin resultante en la molécula, esta emisión de radio frecuencia es detectada por la bobina que genera la radiofrecuencia de excitación (Keeler, 2002).

El tiempo entre la inversión de las poblaciones es denotado por (T_1). Cuando es aplicado un segundo RFP este genera una precesión en fase de los espines magnéticos, al retirar el segundo pulso de RFP, los espines se desfasan debido a la diferencia en ambientes químicos de cada átomo dentro de la molécula (T_2). Las diferencias en relajación generan unas curvas de decaimiento de inducción libre (FID) que dependen del entorno químico del núcleo en la molécula (Duer, 2002).

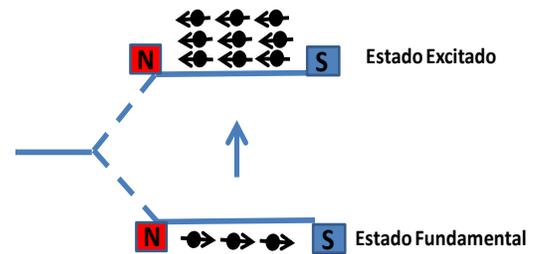
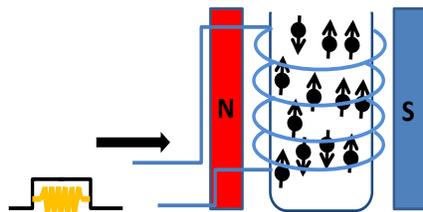
a) Muestra de Polímero Protones en la Cadena



b) Campo Magnético Externo Permanente



c) Pulso de Radio Frecuencia



d) Relajación

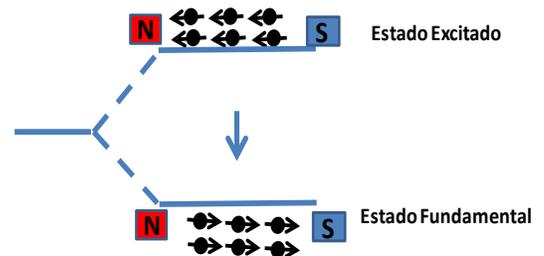
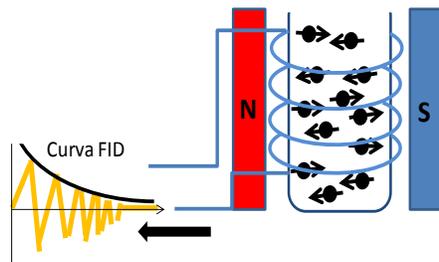


Figura 10. Representación de los vectores resultantes en un proceso de precesión magnética para spines expuestos a un campo magnético externo propios de un proceso de ^1H -RMN durante el análisis del polipropileno.

Las curvas FID son un registro del proceso de decaimiento magnético originado por la ^1H -RMN. El registro obtenido en las curvas FID está en un dominio de tiempo. Tradicionalmente las curvas FID en dominio tiempo son transformadas a un dominio de frecuencia usando la transformada de Fourier (US patent 5 302 897). La transformación de curvas FID de dominio del tiempo a dominio de frecuencias da origen a los espectros ^1H -RMN. En el estudio de partículas de polipropileno son usadas las curvas FID originadas mediante ^1H -RMN con dominio de tiempo (Packer, 1993).

El proceso FID con dominio en el tiempo es modelado mediante la suma de varias funciones matemáticas de naturaleza exponencial (Packer, 1993). En los análisis de ^1H -RMN muchos de los procesos dinámicos son de naturaleza exponenciales (Derbyshire et al.; 2004).

La función exponencial $Y(t) = Ae^{-at}$ con base e , donde t es el tiempo, A es la amplitud y a es el coeficiente de decaimiento, es una función típica de las empleadas en el modelado de curvas FID (Derbyshire, et al.; 2004).

Un ejemplo del uso de la función $Y(t) = Ae^{-ax}$, es el estudio de las señales de decaimiento magnético inducido (decaimiento exponencial) en función del tiempo. Es esencial comprender la naturaleza de las curvas exponenciales, para poder describir y modelar de manera adecuada el fenómeno del decaimiento magnético inducido (FID).

La aplicación de diversas funciones matemáticas al modelado de las curvas FID antes de someterlas a la transformación de Fourier (FT), puede aumentar drásticamente la calidad del espectro obtenido (McKenzie, et al.; 2011). Sin embargo, en general hay un equilibrio entre la resolución de los picos y la relación señal-ruido (SNR).

En el caso de muestras de polipropileno homopolímero y copolímero, existen tres ecuaciones que modelan o describen cada parte del proceso FID para ^1H -RMN: La función Gaussiana rápida descrita por la ecuación 2, La Gaussiana lenta ecuación 3 y función

exponencial, ecuación 4. La Gaussiana rápida es una función de decaimiento asociada a la región cristalina (inmóvil) del polipropileno (US patent 5,302,897). La Gaussiana lenta es una función de decaimiento asociada a la región semi-cristalina (transición cristalina/amorfa) de las partículas de polipropileno y la tercera función es la función exponencial, que describe el proceso FID en la región amorfa (móvil) de la partícula sólida de polipropileno (ecuación 4). La Figura 11 representa las tres curvas que modelan el proceso FID para muestras de polipropileno.

$$Y_{FG} = Ae^{-(at^2)} \text{ (EC. 2)}$$

$$Y_{SG} = Be^{-(bt^2)} \text{ (EC. 3)}$$

$$Y_{Exp} = Ce^{-(ct)} \text{ (EC. 4)}$$

La ecuación general que describe los tres fenómenos se puede escribir de la siguiente forma:

$$f(t) = Ae^{-(at^2)} + Be^{-(bt^2)} + Ce^{-(ct)} \text{ (EC. 5)}$$

Cada una de las 3 curvas que describen el decaimiento FID es generada mediante un proceso interactivo de modelado computacional llamado Marquardt-Levenberg (ML) (US patent 5, 302,897). La técnica de modelado computacional M-L, origina los valores de amplitud y los valores de índice de decaimiento para cada una de las ecuaciones exponenciales que modelan las curvas FID de polipropileno. Luego de generados los 6 parámetros matemáticos (amplitud e índice de decaimiento para cada curva) son conmutados para obtener 230 descriptores y generar un modelo de predicción de la propiedad química o física de interés. La interpretación de las curvas FID y de los descriptores matemáticos generados requiere la aplicación de métodos de análisis multivariado. El análisis multivariado es útil en interpretación y generación de modelos de regresión que predigan las propiedades químicas o físicas de interés en las muestras de polipropileno producidas.

El uso tradicional de las técnicas espectroscópicas aplicadas a procesos de producción ha sido centrado en identificación estructural y de grupos funcionales en diferentes tipos de sustancias compuestos químicos. La aplicación de espectroscopia Raman, Infrarrojo cercano (NIR), infrarrojo medio (MID) y de ^1H -RMN fue extendida posteriormente a la determinación cuantitativa de los componentes en una muestra relacionando la intensidad de la señal espectral de un grupo funcional con la concentración.

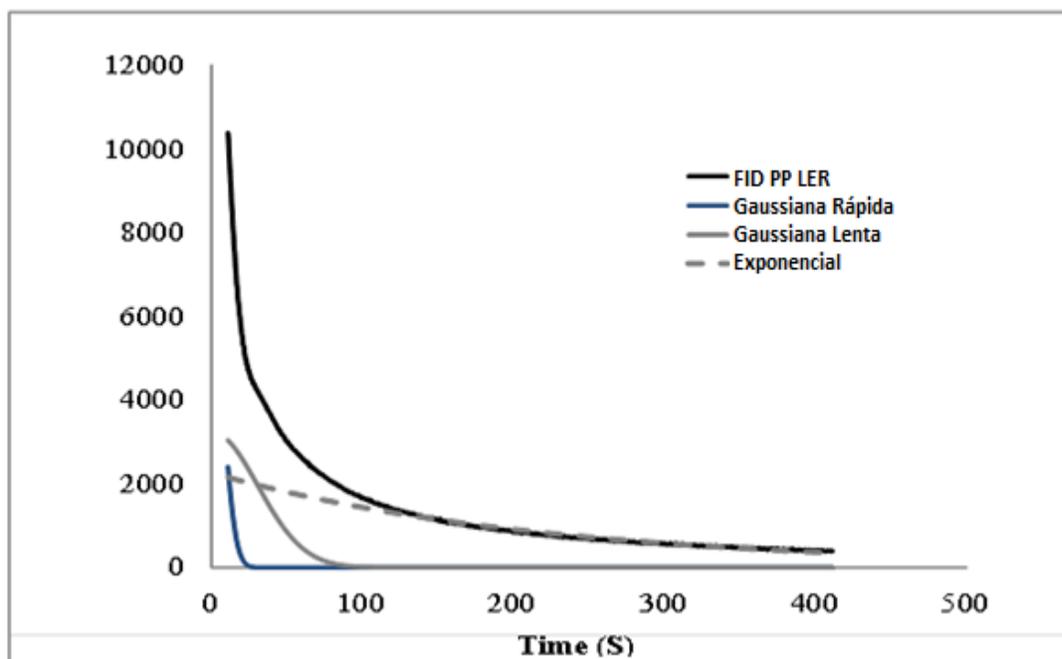


Figura 11. Representación gráfica de la curva de decaimiento FID y de las tres curvas que modelan el proceso de decaimiento magnético.

La digitalización de las curvas de FID del ^1H -RMN permitió relacionar las características espectrales con las propiedades físicas, químicas y mecánicas por medio de herramientas estadísticas y matemáticas (quimiometría). Por ejemplo P.E. Kristiansen, et al.:(2000), estudiaron los cambios en la señal protónica de ^1H -RMN para decaimiento libre de inducción magnética de una muestra de polietileno mientras se calientan de 360 a 400 K. La cristalinidad, según lo determinado por el ajuste del modelo, disminuye de forma lineal con el aumento de temperatura desde los 360 hasta 393 K. A temperaturas mayores a los 393 K, la cristalinidad disminuye rápidamente (de forma no lineal) al aumentar la

temperatura. Dentro de la región de temperaturas estudiada, tres diferentes fases son identificadas. El objetivo principal de este trabajo fue hacer el uso potencial de la técnica de análisis de ^1H -RMN mediante la curva FID para caracterizar polietileno (PE) y otras poliolefinas. El estudio presenta una función matemática que describe la curva FID ^1H -RMN para la fase cristalina del PE sólido. Dos curvas exponenciales describen la porción amorfa de la curva FID de RMN del PE sólido y una tercera ecuación matemática que describe la región de transición cristalina/amorfa en la curva FID ^1H -RMN del PE sólido. Los resultados muestran que para lograr una correcta descripción del fenómeno es necesario combinar varias expresiones matemáticas (Función Pake, Función Weibulian y Función exponencial; Función exponencial o Función Abragamian-Weibulian y Función exponencia) teniendo en cuenta los diferentes rangos de temperatura (Kristiansen, Hansen, y Pedersen, 2000).

Lili Zhang, et al.; (2009) describen la cristalinidad de copolímero de eteno-1-hexeno usando ^1H -RMN y $^{13}\text{C}_6$ -RMN. El estudio realiza una comparación entre los resultados obtenidos mediante las resonancias ^1H -RMN y $^{13}\text{C}_6$ -RMN. Dentro de sus resultados Zhang discute la función Abragamian y la función Pake para estudiar la cristalinidad del PE usando ^1H -RMN. Los resultados muestran concordancia entre la cristalinidad determinada por las curvas FID de ^1H -RMN y los resultados obtenidos con $^{13}\text{C}_6$ -RMN y calorimetría diferencial de barrido (L. Zhang, et al.; 2009).

Hans Karlstrom, et al.;(1995), emplearon el análisis de componente principal para estudiar espectros $^{13}\text{C}_6$ -RMN en muestras sólidas de origen orgánico. Los espectros $^{13}\text{C}_6$ -RMN antes y luego de aplicarles transformada de Fourier fueron usados en la medición del efecto de diferentes condiciones ambientales en la producción de turba a partir de plantas. Los resultados indican que el análisis de componente principal extrae la información contenida en los espectros y las curvas FID de la $^{13}\text{C}_6$ -RMN. Uno de los resultados más importantes es la extracción de la información directamente del interferograma del proceso de decaimiento magnético. En su análisis de la información obtenida Karlstrom describe la información de las muestras condensada en los scores del PCA, y la información obtenida

en los loading del PCA. En sus conclusiones recomienda evitar la normalización de espectros de $^{13}\text{C}_6$ -RMN obtenidos de muestras orgánicas de origen vegetal. También recomienda profundizar el estudio de espectros FID antes de aplicar transformada de Fourier (Karlstrom, 1995).

Hansen, et al.:(1998), resolvieron la señal de espectroscopia ^1H -RMN en estado sólido para el decaimiento de inducción libre (FID). La señal FID del PE fue resuelta en una parte cristalina y una parte amorfa que permite el ajuste del modelo. La parte "cristalina" del FID está representada por la transformada inversa de Fourier de la función "Pake" (en el dominio de la frecuencia). Del mismo modo, la parte "amorfa" de la señal FID es modelada como una combinación lineal de las funciones analíticas de los siguientes tipos: Exponencial (E), Gauss (G), Weibullian (W), y Brereton (B). Los Modelos de ajuste y evaluación estadística de los diferentes conjuntos de funciones, sugieren que el modelo de PBE (funciones Pake, Brereton y Exponencial) permite la mejor representación de la curva FID del polietileno. La cristalinidad es estimada como un coeficiente entre la intensidad de la señal de la función de Pike y la señal de la función de Abragamian en el tiempo cero. La cristalinidad derivada de estos ajustes del modelo se analiza y se compara con los métodos convencionales de cristalinidad determinantes, como las mediciones de densidad y calorimetría diferencial de barrido (DSC). Los resultados obtenidos muestran buena correlación entre la cristalinidad reportada por DSC, la cristalinidad calculada mediante la densidad y la obtenida usando parámetros FID (Hansen, et al.; 1998).

Litvinoov (2006), determino la composición estructural de gomas de aceite extendido etileno-propileno-dieno (EPDM) y polipropileno, EPDM y aceite termo-vulcanizados (TPVs) usando RMN de estado sólido. Los resultados demostraron que el método RMN de estado sólido permite la caracterización selectiva de las fases cristalinas y amorfas de PP, reticulado EPDM, y el aceite en TPVs. El método predice la densidad de red en gomas de polipropileno y EPDM luego de TPVs. El método de análisis mide la composición de fase del material TPVs y cuantifica la movilidad de las moléculas de aceite en el material estudiado (Litvinov, 2006).

Pandey, et al.;(2002), realizaron una evaluación no destructiva de las propiedades físico-mecánicas de polipropilenos copolímeros de impacto usando espectroscopia infrarroja. En su estudio reportaron el desarrollo de un método alternativo, rápido directo y sencillo para determinar propiedades que son típicamente requeridas en la industria para categorizar la calidad de los polipropilenos copolímeros de impacto. Encontraron a través de técnicas de análisis multivariantes una buena concordancia entre los resultados obtenidos con el modelo y los resultados obtenidos mediante los ensayos tradicionales para determinar las propiedades como Melt Flow Index (MFI); Impacto Izod y Módulo de flexión, en sus resultados indican que la composición del polipropileno copolímero determina las propiedades mecánicas del material estudiado (Pandey, 2002).

Sardashti y Lai (2001), mediante la patente número US 6,208,137 B1; presentaron un método quimiométrico para la predicción de propiedades de polímeros con un sistema RMN on-line. En esta reclaman la invención de una metodología que inicia con la obtención de un conjunto de datos que relaciona las curvas de decaimiento de inducción libre (FID) con los valores medidos de la propiedad Solubles en Xileno de muestras de polipropileno. Este conjunto de datos se analiza usando el análisis de componentes principal (PCA) con lo cual se obtiene un conjunto de componentes principales. Entonces los tres conjuntos: el de los componentes principales, el de los solubles en xileno del polipropileno y el de los FID se analizan mediante un análisis mínimos cuadrados parciales para obtener un conjunto de datos de entrenamiento para ser validados posteriormente para obtener un conjunto de datos predictivos. A partir de los datos predictivos al obtener los FID de muestras con valores desconocidos de Solubles en xileno se puede predecir su valor.

La ventajosa capacidad de los métodos quimiométricos para predecir la composición de muestras a partir de datos espectroscópicos fue presentada por Salomonsen, et al.; (2008), demostraron que es posible desarrollar un método cuantitativo robusto no destructivo y rápido para la medición de monómeros de polvos de alginato comercial usando

espectroscopia vibracional (IR, Raman y RMN) y métodos quimiométricos. Lo sobresaliente de este logro radica en que las técnicas espectroscópicas poseen un gran potencial de aplicación industrial para el control de calidad en línea así como para el screening de un gran número de muestras. El modelo con Raman es simple y obtenido con un solo componente PLS con un error de predicción de 0,08 (raíz cuadrada media del error de predicción) y un coeficiente de correlación (R) de 0,97 que es comparable con la desviación estándar del método de referencia.

2.7 QUIMIOMETRÍA

La quimiometría se ha definido de diversas formas por varios autores como “El arte de extraer información química relevante de los datos producidos por un experimento químico” (Wold, 1995). Sin embargo, más recientemente se ha definido de manera más amplia como “La disciplina que usa la matemática, estadística y la lógica formal para:

- a) Diseñar o seleccionar procedimientos experimentos óptimos
- b) Proveer la máxima información química relevante para analizar los datos químicos;
- c) Obtener conocimiento acerca del sistema químico” (Massart, et al.; 1997).

Las curvas FID generadas en la espectroscopia de ^1H -RMN carecen de picos de desplazamiento debido a la absorción (McKenzie, et al.; 2011). La ausencia de estas en las curvas FID, dificulta obtener información química o física de las propiedades del polímero al realizar un análisis simple por inspección. Para interpretar y extraer la información contenida en las curvas FID de ^1H -RMN es necesario realizar un análisis multivariado. Por lo tanto, la aplicación de este tipo de análisis a los datos de química analítica se denomina quimiometría (McKenzie, et al.; 2011, Buydens, 1999).

Los métodos quimiométricos usados en las curvas FID de ^1H -RMN proporcionan una serie de ventajas ausentes en el análisis univariado (Karlstrom, 1995). La aplicación de la quimiometría aplicada proporciona la máxima información química relevante mediante el análisis de los datos químicos, para obtener conocimientos sobre los sistemas químicos

(Bro, 2003). Los análisis quimiométricos pueden ser usados para conocer parámetros de producción que permiten lograr el control y asegurar la calidad final de un proceso industrial o un producto terminado (Cabri, et al.; 2007).

Existen en quimiometría dos grandes divisiones según la aplicación, análisis cualitativo y análisis cuantitativo. El análisis cualitativo de la información permite observar tendencias, agrupamientos o aleatoriedad en los datos obtenidos, para ello se puede usar el análisis por componentes principales (PCA). El análisis cuantitativo de la información química, permite medir, calcular o predecir una o varias propiedades de las muestras de interés, como en el caso de la regresión por mínimos cuadrados parciales (PLSR), regresión lineal múltiple (MLR) o la regresión de componentes principales (PCR). Cada análisis quimiométrico cualitativo o cuantitativo provee información relevante al momento de interpretar información química propias de una reacción o un proceso de producción (Wise, 1996).

2.7.1 Pre-tratamientos de los Datos

Cuando los datos son obtenidos directamente de un instrumento, estos salen con ruido o alteraciones inevitables que pueden afectar la calidad del análisis multivariado, por lo cual es necesario realizar algunas transformaciones o pre-tratamientos matemáticos para mitigar o reducir algún efecto sobre los datos.

Las transformaciones matemáticas, previo al análisis de datos multivariados, ayudan a reducirlos efectos del ruido aleatorio, cambios en la línea base o diferencias en las variables por la escala de estas. Los pre-tratamientos de datos son aplicados antes del análisis multivariado con el fin de reducir, eliminar o uniformar el impacto de variables indeseadas en los datos adquiridos. Una cuidadosa selección de los pre-tratamientos de datos puede mejorar significativamente la robustez de un modelo de calibración, por lo tanto son usados de acuerdo a los efectos que son capaces de corregir en los datos crudos o como son obtenidos del instrumento (Bakeev, 2005).

El pre-tratamiento más simple comúnmente usado es la corrección de línea base, el cual consiste en sustraer a cada curva FID, la media de todas las curvas FID. Los pre-tratamientos matemáticos utilizados para compensar los desplazamientos de línea base incluyen la corrección de dispersión multiplicativa (MSC) y la corrección por variación normal estandarizada (SNV). Estos pre-tratamientos, SNV y MSC, fueron desarrollados originalmente para procesar las curvas FID de reflexión de sustancias sólidas, pero luego fueron también aplicados a curvas FID de transmisión en sólidos (Reich, 2005). Los cambios en línea base y las diferencias de intensidad espectral pueden ser reducidos o eliminados por los algoritmos de normalización (Bakeev, 2005).

Por otro lado, la aplicación de la primera y la segunda derivada pueden mejorar la resolución de las bandas superpuestas. Además, son capaces de reducir las diferencias en línea base. Teniendo en cuenta que el ruido espectral es amplificado con la aplicación de las derivadas, las generalmente, las derivadas se combinan con métodos de suavizado tipo Taylor o Savitzky-Golay (Reich, 2005).

2.7.2 Análisis Multivariado de Datos

Los datos en el análisis multivariado se refieren a:

- **Objetos:** Son muestras, moléculas químicas o curvas FID. Son descritas en la matriz X como n filas de objetos y típicamente van desde de 20 a 1000 objetos.
- **Variables características:** Llamadas también variables descriptivas o simplemente descriptores, los cuales son medidas o calculadas. En la matriz X son las m columnas de descriptores, típicamente son de 3 o más variables descriptivas. Ver una representación gráfica en la figura 12.

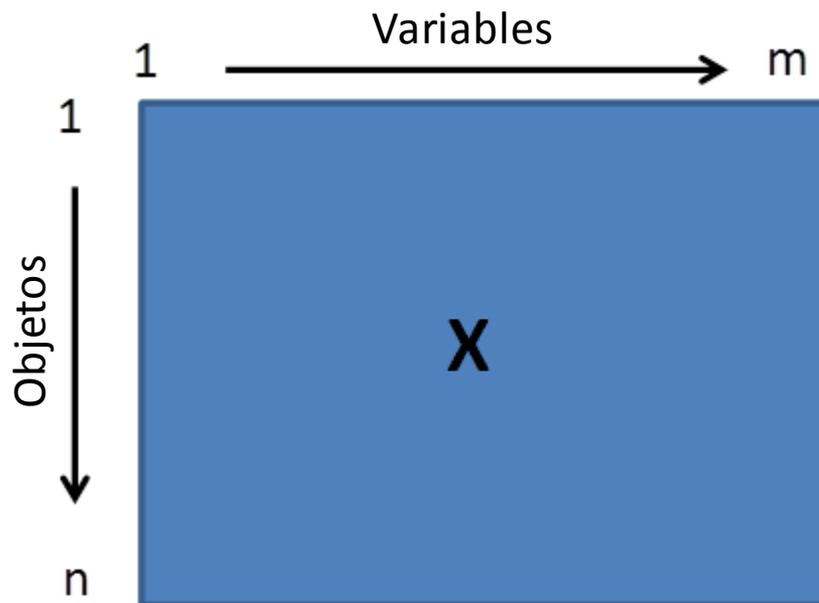


Figura 12. Matriz X de n filas y n columnas.

2.7.2.1 Análisis Multivariado Cualitativo

En el análisis cualitativo, las variaciones espectrales de las muestras son relacionadas con propiedades de las muestras que poseen valores discretos, que representan la identidad del producto o que indican la calidad del producto, por ejemplo bueno o malo (Reich, 2005) los métodos multivariados de análisis cualitativo son empleados para agrupar muestras con características similares (Mertens, et al.; 1994). Los métodos de análisis cualitativo están subdivididos en los algoritmos de aprendizaje supervisado y no supervisado, dependiendo de si la clase a la que pertenecen las muestras se conoce (Reich, 2005). Es decir si los valores de las propiedades a medir se conocen se dice que son métodos supervisados.

Métodos no supervisados, también conocidos como análisis de conglomerados, no requieren un conocimiento a priori sobre la estructura del grupo de datos. Los métodos de análisis de conglomerados generan agrupación, es decir, clasifican las diferentes muestras en subgrupos que permite la aglomeración en sí misma. Este tipo de análisis suele ser muy

útil en las primeras etapas de una investigación para explorar las sub-poblaciones en un conjunto de datos, por ejemplo, diferentes materiales con diferentes grados de propiedades físicas. El análisis de conglomerados se puede realizar con simples técnicas visuales, como el análisis de componentes principales (PCA) o algunos de los métodos jerárquicos que conducen a los gráficos llamados dendrogramas.

La mayoría de los métodos de análisis cualitativo se basan en medidas de la similitud (Candolfi, et al.; 1999). La similitud (S) entre objetos es expresada en términos de los coeficientes de correlación o distancia (Reich, 2005).

Métodos de clasificación supervisada, también conocidos como análisis discriminante, se utilizan para construir las reglas de clasificación para una serie de subgrupos pre-especificados, es decir, la estructura del grupo del conjunto de entrenamiento es conocido (Bakeev, 2005). Las reglas de clasificación construidas con el análisis supervisado, son utilizadas para la asignación de muestras desconocidas al subgrupo más probable. La identidad o la calidad de bueno / malo de una muestra son, por tanto, decididas al clasificar la muestra como perteneciente a un grupo de propiedades conocidas.

Algoritmos de clasificación supervisada son el análisis discriminante lineal (LDA), análisis discriminante cuadrático (QDA), modelo suave e Independiente de analogías de clase (SIMCA) o el algoritmo del vecino más cercano (KNN), estos algoritmos son utilizados normalmente para la construcción de bibliotecas espectrales (Reich, 2005). La mayoría de los métodos de clasificación pueden operar en el espacio de longitud de onda o en una dimensión reducida en espacio de factores. En cualquier caso, su objetivo final es establecer criterios matemáticos para la parametrización de similitud espectral, por lo que permite la comparación entre las muestras o de una muestra y un grupo de muestras con propiedades conocidas (Reich, 2005;Bakeev, 2005)

Para lograr la clasificación se requieren amplias bibliotecas de espectros que representan la variación natural de cada producto. Las bibliotecas espectrales tienen que ser construidas en un proceso de calibración, y la similitud entre muestras es expresada ya sea por un

coeficiente de correlación o de distancia Mahalanobis o distancia Euclidiana, como el valor de concordancia espectral (Reich, 2005).

2.7.2.1.1 Reducción de variables, análisis de componentes principales (PCA)

Los métodos instrumentales de $^1\text{H-RMN}$ generan un gran volumen de información, por lo que uno de los aspectos más importantes dentro de la quimiometría es el desarrollo de métodos capaces de reducir este gran volumen de información a un número reducido de variables. Uno de los métodos de reducción de variables más empleados es el Análisis en Componentes Principales (PCA).

PCA es un procedimiento matemático que resuelve los datos espectrales en componentes ortogonales, cuyas combinaciones lineales es aproximada a los datos originales. Las nuevas variables, llamadas componentes principales (PC), que pueden ser los loadings o los scores, corresponden a los mayores eigen valores o valores propios de la matriz de covarianza, por lo tanto, lo que representa la mayor variación posible en el conjunto de datos. El primer PC representa la variación máxima en la matriz de covarianza, correspondiente con la mayor fuente de variación dentro de las muestras estudiadas (Shlens, 2005).

El método PCA es descrito en la ecuación 6. La matriz X representa los datos originales organizados en filas y columnas (curvas FID o descriptores matemáticos), T es la matriz de scores, con la variación de las muestras comprimida, P' es la transpuesta de P que representa a la matriz que contiene los loadings o vectores de variación de variables. La matriz E representa el error o variación no explicada por el PCA.

El número de componentes relevantes, a , permite determinar la matriz de scores Ta altamente correlacionada con la matriz X , donde los a PCA scores son no correlacionados entre sí. La ecuación 7 contiene la matriz residual E . El límite para que un score sea relevante es cuando su varianza es mayor del 1% de la varianza total.

$$X = TP' \quad (\text{EC. 6})$$

$$X = T_a P'_a + E \quad (\text{EC. 7})$$

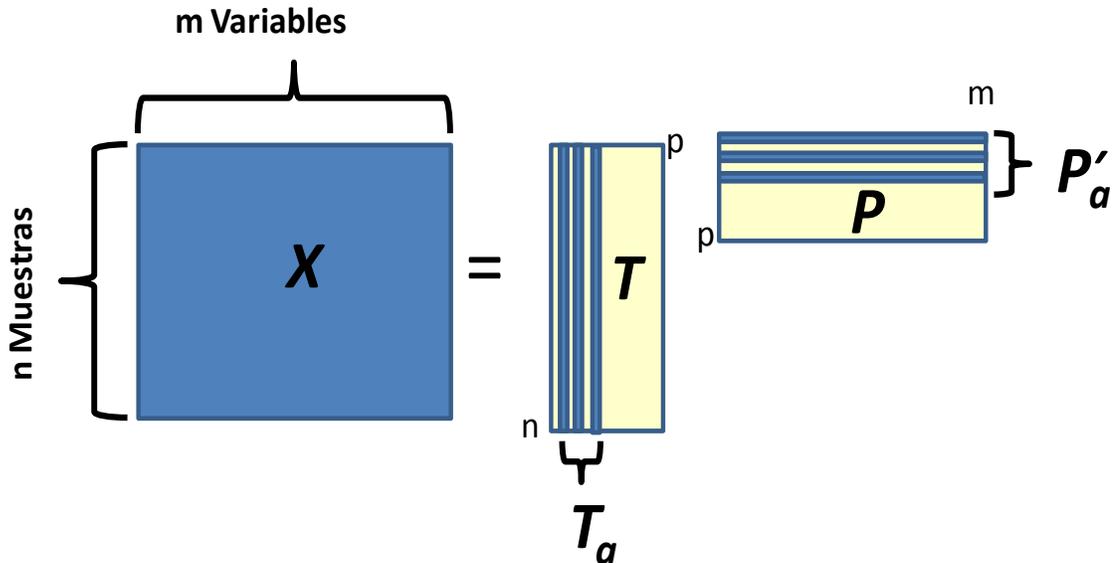


Figura 13. Representación gráfica de las matrices envueltas en un análisis de componente principal.

2.7.2.1.2 Componentes Principales

Uno de los grandes propósitos del análisis multivariado es encontrar nuevas variables denominadas variables latentes o escores expresadas en función de las variables descriptivas. Los escores son por tanto una combinación lineal de las variables originales, las cuales condensan en forma apropiada en pocas variables la información que contienen las variables descriptivas. Típicamente, las variables descriptivas son un número grande que sobre pasan cientos de estas y en algunos casos miles, y son reducidas a 2 o 3 escores, los cuales son correlacionadas con las propiedades de los objetos. Dependiendo del propósito del método multivariado se utiliza diferente criterio matemático para obtener las

variables latentes. En el caso de componentes principales, la matriz de datos X o variables aleatorias X_1, X_2, \dots, X_p

a X una matriz de vectores columnas de variables aleatorias con una matriz de varianza covarianza S_x .

$$X_{n \times p} = \begin{matrix} & X_1 & X_2 & \dots & X_j & \dots & X_p \\ \begin{matrix} X_{11} & X_{12} & \dots & X_{1j} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2j} & \dots & X_{2p} \\ X_{31} & X_{32} & \dots & X_{3j} & \dots & X_{3p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nj} & \dots & X_{np} \end{matrix} \end{matrix} \quad (\text{EC. 8})$$

Donde $X_j = x_{ij} - x_j \quad j= 1, 2, \dots, p$

Para los cual es posible obtener una matriz de $V_{p \times p}$ que permite transformar la matriz de datos $X_{n \times p}$ en una nueva matriz $U_{n \times p}$ en la cual cada fila representa una combinación lineal de las variables originales.

$$V_{p \times p} = \begin{matrix} & v_1 & v_2 & \dots & v_j & \dots & v_p \\ \begin{matrix} v_{11} & v_{12} & \dots & v_{1j} & \dots & v_{1p} \\ v_{21} & v_{22} & \dots & v_{2j} & \dots & v_{2p} \\ v_{31} & v_{32} & \dots & v_{3j} & \dots & v_{3p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ v_{p1} & v_{p2} & \dots & v_{pj} & \dots & v_{pp} \end{matrix} \end{matrix} \quad (\text{EC. 9})$$

$$U_{n \times p} = X_{n \times p} \cdot V_{p \times p} \quad (\text{EC. 10})$$

Con una combinación lineal de

$$u_1 = X \cdot v_1 = x_{11}v_{11} + x_{12}v_{21} + \dots + x_{1i}v_{i1} + \dots + x_{1p}v_{p1} \quad (\text{EC. 11})$$

$$u_2 = X \cdot v_2 = x_{21}v_{12} + x_{22}v_{22} + \dots + x_{2i}v_{i2} + \dots + x_{2p}v_{p2}$$

...

...

$$u_n = X \cdot v_1 = x_{n1}v_{1p} + x_{n2}v_{2p} + \dots + x_{ni}v_{ip} + \dots + x_{np}v_{pp}$$

Entonces

$$Var(u_i) = v_i S_x v_i \quad i=1,2,\dots,p$$

$$Cov(u_j, u_k) = v_j S_x v_k, \quad k=1,2,\dots,p$$

Las combinaciones lineales obtenidas son conocidas como componentes principales o brevemente como CP's: u_1, u_2, \dots, u_n , cuyas varianzas son tan grande como es posible. El primer componente contiene la primera combinación lineal con la máxima varianza y así en general para el resto PC's:

$$Var(u_1) > Var(u_2) > \dots > Var(u_i) > \dots > Var(u_n) \quad (EC. 12)$$

Para que se cumple la condición de independencia entre los componentes principales, los vectores propios de la matriz de escores su longitud debe ser igual a la unidad, además:

$$PrimerCP = \sum_{i=1}^p x_{1i} \cdot v_{i1} \quad (EC. 13)$$

Su varianza $Var u_1 = Var(X_{1i} \cdot v_{i1})$ es máxima sujeto a $v_1^T \cdot v_1 = 1$

$$SegundoCP = \sum_{i=1}^p x_{2i} \cdot v_{i2} \quad EC. 14$$

Su varianza $Var u_2 = Var(X_{2i} \cdot v_{i2})$ es máxima sujeto a $v_2^T \cdot v_2 = 1$ y $v_1^T \cdot v_2 = 0$

Además se puede probar que $Cov u_1, u_2 = 0$

En el k-ésimo CP

$$k - \text{enesimoCP} = \sum_{i=1}^m x_{ki} \cdot v_{ik} \quad (\text{EC. 15})$$

Su varianza $\text{Var } u_k = \text{Var}(X_{k,i} \cdot v_{i,k})$ es máxima sujeto a $v_k^T \cdot v_k = 1$

$$v_1^T \cdot v_k = 0 \text{ y } v_j^T \cdot v_k = 0 \text{ para } j \neq k \quad (\text{EC. 16})$$

Además se puede probar que $\text{Cov } u_j, u_k = 0$

2.7.2.1.3 Algebra Matricial y Componentes Principales

El método de componentes principales se fundamenta en las propiedades del algebra de matrices, en donde la matriz de covarianzas $S_{p \times p}$ es una matriz simétrica generada a partir de la matriz de datos $X_{n \times p}$.

$$X = [x_1 \ x_2 \ \dots \ x_i \ \dots \ x_p] \quad (\text{EC. 17})$$

$$S_x = \frac{1}{n-1} X - 1X^T X - 1X \quad (\text{EC. 18})$$

$$x_{ki} = x_{ki} - x_i \quad (\text{EC. 19})$$

x_{ki} es un vector centrado con media igual 0 y una desviación estándar S_i

$$S_{p \times p} = \begin{pmatrix} s_1 & s_{12} & \cdots & s_{1j} & \cdots & s_{1p} \\ s_{21} & s_2 & \cdots & s_{2j} & \cdots & s_{2p} \\ s_{31} & s_{32} & \cdots & s_{3j} & \cdots & s_{3p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{p-1,1} & s_{p-1,2} & \cdots & s_{p-1,j} & \cdots & s_{p-1,p} \\ s_{p1} & s_{p2} & \cdots & s_{pj} & \cdots & s_p \end{pmatrix} \quad (\text{EC. 20})$$

$$s_i^2 = \text{var } X_i = \frac{SS_i}{n-1} = \frac{x_i - \bar{x}}{n-1}^2 \quad (\text{EC. 21})$$

$$s_{ij}^2 = \text{cov } X_i, X_j = \frac{SS_{ij}}{n-1} = \frac{x_i - \bar{x}_i}{n-1} \frac{x_j - \bar{x}_j}{n-1} \quad (\text{EC. 22})$$

La matriz S de dimensión $p \times p$

$$R_{p \times p} = \begin{pmatrix} 1 & r_{12} & \cdots & r_{1j} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2j} & \cdots & r_{2p} \\ r_{31} & r_{33} & \cdots & r_{3j} & \cdots & r_{3p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pj} & \cdots & 1 \end{pmatrix} \quad (\text{EC. 23})$$

$$r_{jk} = \frac{\text{cov}(X_j, X_k)}{\sqrt{\text{var}(X_j)\text{var}(X_k)}} \quad (\text{EC. 24})$$

2.7.2.1.4 Cálculo de Matriz de Loading (V): Método de Eigenvalores y Eigenvectores

De acuerdo a las propiedades del algebra de matrices la matriz de varianza y covarianza puede ser reducida a una matriz diagonal D al multiplicar tanto por la izquierda como por la derecha por la matriz ortogonal V tal que

$$D = V^T \cdot S \cdot V \quad (\text{EC. 25})$$

Los elementos diagonales de D , $\lambda_1, \lambda_2, \dots, \lambda_p$ son llamadas raíces características o valores propios de S . Las columnas de V , v_1, v_2, \dots, v_p son llamada vectores característicos o vectores propios de S . Las raíces características pueden ser obtenidas de la solución de la ecuación determinante, llamada ecuación secular o característica.

$$Sv = \lambda v \text{ (EC. 26)}$$

Lo cual es equivalente a

$$Sv = \lambda Iv \text{ (EC. 27)}$$

Reordenado esto se tiene

$$S - I\lambda \quad v = 0 \text{ (EC. 28)}$$

Si S es una matriz $p \times p$, la ecuación corresponde a un sistema homogéneo de p ecuaciones con incógnita u . Se supone que el sistema tiene soluciones no triviales, para lo cual $\det(S - \lambda I) = 0$. Donde λ es el valor propio de S y I es la matriz identidad.

2.7.2.1.5 Calculo de los Componentes Principales: Matriz de Scores (U)

La transformación obtenida de los ejes principales debe transformar las p correlacionadas variables x_1, x_2, \dots, x_p en p nuevas variable no correlacionadas u_1, u_2, \dots, u_p . Los nuevos ejes coordenados de estas nuevas variables son descritos por vectores característicos v_i con la cual se construye la matriz u de dirección de cosenos usado en la transformación.

$$u = v^T x^T - Ix \text{ (EC. 29)}$$

Aquí x y $I.x$ son vectores de dimensión $p \times 1$ de las observaciones en relación a la variable original.

Las variables transformadas son llamadas componentes principales de x o abreviado como CP.

El componente principal i -enésimo es:

$$v_i = v_i^T x_i^T - I x_i \quad (\text{EC. 30})$$

Consiste en que para el cual tenemos una media de cero y una varianza λ_i igual al i -enésimo raíz del valor propio.

Una distinción entre variables transformadas y observaciones transformadas consiste en que las variables transformadas son llamadas componentes principales y las n variaciones son llamadas scores.

2.7.2.1.6 Cálculo de la matriz de loading (V): Método de SDV (Singular Value Decomposition)

$$X_{n \times p} = U_{n \times n} D_{n \times p} V_{p \times p}^T \quad (\text{EC. 31})$$

Donde U y V son ortogonales

$$U^T U = I_{n \times n} \quad (\text{EC. 32})$$

$$V^T V = I_{p \times p} \quad (\text{EC. 33})$$

$$\text{diag}(D) = \lambda_1, \lambda_2, \dots, \lambda_n \quad (\text{EC. 34})$$

El cálculo de SDV consiste en encontrar los eigenvalores ($\text{diag}(D)$) y eigenvectores (v_i) de $X^T X$ o XX^T . Con los eigenvectores de $X^T X$ se construye la matriz V y la matriz U .

2.7.2.2 Análisis Multivariado Cuantitativo

Previo a cualquier análisis cuantitativo, es requerido entrenar el modelo de predicción, es decir, calibrarlo usando métodos multivariados. El proceso de calibración consiste básicamente en los siguientes pasos:

1. La selección de un conjunto representativo de la muestra de calibración.
2. Adquisición de curvas de FID y determinación de los valores de referencia.
3. Modelado multivariado para describir la variación espectral respecto a los valores de referencia de la propiedad de destino de análisis.
4. Validación del modelo de predicción, usando validación cruzada, o la validación externa.

Los métodos de regresión multivariada más empleados son la regresión por componentes principales (PCR) y la regresión por mínimos cuadrados parciales (PLS) (Wold, 2001).

PCR utiliza los componentes principales encontrados en el PCA para realizar la regresión de la propiedad muestra que se desea predecir. PLS encuentra las direcciones de mayor variabilidad al comparar la variación espectral y la información de propiedad, con esta comparación crea unos nuevos ejes, denominados componentes o factores PLS (Reich, 2005). La principal diferencia entre los métodos PLS y PCR es que el primer componente principal o factor de PCR representa las mayores variaciones en la curva de FID mientras que en PLS que representa las variaciones más relevantes que muestra la mejor correlación con los valores de propiedad de destino. En ambos casos, el número óptimo de factores que se utilizan para construir el modelo de calibración depende de las propiedades de la muestra y el objetivo de análisis (Reich, 2005).

La selección del número de factores requeridos por el modelo de calibración es importante para el desempeño del modelo de predicción. Seleccionar un número de factores excesivo puede conducir a un modelo sobrealimentado. Un modelo de predicción sobrealimentado presenta coeficiente de regresión cercano a 1.0, bajo error de predicción para las muestras de calibración, pero valores altos de error de predicción para muestras de validación (muestras externas al grupo de calibración).

Un modelo sobrealimentado no es muy robusto y puede fallar cuando se prueba con muestras de validación independientes (Reich, 2005). A continuación se describen los métodos de análisis cuantitativo empezando con los modelos univariados hasta abarcar la regresión PLS.

2.7.2.2.1 Regresión Lineal Simple

Sea

$$y_i = \beta_0 + \beta_1 x_i \quad (\text{EC. 35})$$

Donde β_0 =Intercepto, β_1 = Pendiente

- Asunciones Estadísticas

a) Existencia:

Sea una pareja X y Y , donde

X = Variable independiente, fija y no aleatoria.

Y = Variable dependiente. Y es una variable aleatoria con cierta probabilidad de distribución con una media y una varianza finita.

Definimos

$\mu_{Y/X}$ = Media poblacional, donde la variable Y depende de X

$\sigma_{Y/X}$ = Varianza poblacional, donde la variable Y depende de X

b) Independencia: Los valores de Y son estadísticamente independientes unos de otros. Cuando no hay independencia entre los valores de Y puede conducir a conclusiones estadísticas invalidas.

c) Linealidad: La media de Y, $\mu_{Y/X}$, es una función lineal de X.

Lo cual puede ser expresado con la ecuación siguiente:

$$\mu_{Y/X} = \beta_0 + \beta_1 X \quad (\text{EC. 36})$$

Por lo tanto, la ecuación de regresión sería:

$$Y = \beta_0 + \beta_1 X + E \quad (\text{EC. 37})$$

Donde E es el error de los valores esperados. E es una variable aleatoria con una media de 0 , $\mu_{E/X} = 0$, para un X cualquiera.

Donde el error se puede despejar de esta forma:

$$E = Y - \beta_0 + \beta_1 X = \mu_{Y/X} - (\beta_0 + \beta_1 X) \quad (\text{EC. 38})$$

d) Homocedasticidad: Homo = el mismo; scedastic = aleatorio, donde la varianza de Y es la misma para cualquier X. Cuando no se aplica esta asunción se denomina, heterocedasticidad. Esto significa que la varianza de Y, $X_1 = \sigma^2_{Y/X1}$, es mayor que la varianza de Y, $X_2 = \sigma^2_{Y/X1}$.

Dados estos términos la homocedasticidad es:

$$\sigma^2_{Y/X} = \sigma^2 \quad (\text{EC. 39})$$

o también como

$$\sigma^2_{Y/X_i} = \sigma^2_{Y/X_j} \text{ (EC. 40)}$$

e) Distribución Normal: Para cualquier valor fijo de X, Y tiene una distribución normal. Esta asunción es posible evaluar mediante las pruebas de significancia estadística (Ej. intervalos de confianza de la media y pruebas de hipótesis).

Hay una clara diferencia entre variable aleatoria, parámetro y estimación puntual.

Y es una variable aleatoria y las observaciones Y_i producen un valor particular de la variable fija X_i (variable conocida).

β_0 y β_1 son parámetros con valores desconocidos pero específicos para una población particular.

β_0 y β_1 , son puntos de estimación de β_0 y β_1 , El estimador E es denominado residual. Para n parejas de $X, Y (x_1, y_1, \dots, x_i, y_i, \dots, x_n, y_n)$, hay n residuales $E_i (E_1, \dots, E_n)$, para $i = 1, 2, \dots, n$.

2.7.2.2.2 Método de Mínimos Cuadrados

Los modelos se construyen con valores observados para X y Y. En la cual se estimen los parámetros de β_0 y β_1 , de tal forma que los valores de X y Y se ajustan a la mejor función lineal que minimice la suma de los valores residuales al cuadrado.

$$e_i^2 = (Y - \hat{Y})^2 = [Y - \beta_0 + \beta_1 X]^2 \quad \text{(EC. 41)}$$

Los mínimos cuadrados determinan los valores de β_0 y β_1 o el mejor polinomio lineal que minimice la suma del cuadrado de los residuales (L), tal que:

$$L = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \beta_0 + \beta_1 x_i)^2 \quad \text{(EC. 42)}$$

Se obtiene el valor mínimo al derivar parcialmente L respecto de β_0 y β_1 .

$$\frac{\partial L}{\partial \beta_0} = \sum_{i=1}^n 2(y_i - \beta_0 - \beta_1 x_i) = 0 \quad (\text{EC. 43})$$

$$\frac{\partial L}{\partial \beta_1} = \sum_{i=1}^n 2(y_i - \beta_0 - \beta_1 x_i) \cdot x_i = 0 \quad (\text{EC. 44})$$

Reordenado los términos

$$\sum_{i=1}^n \beta_0 + \beta_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \quad (\text{EC. 45})$$

$$\sum_{i=1}^n \beta_0 x_i + \beta_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \quad (\text{EC. 46})$$

Simplificando las expresiones

$$n\beta_0 + \beta_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \quad (\text{EC. 47})$$

$$\beta_0 \sum_{i=1}^n x_i + \beta_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \quad (\text{EC. 48})$$

En forma matricial

$$\begin{bmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \end{bmatrix} \quad (\text{EC. 49})$$

Resolviendo el sistema de ecuaciones:

$$\beta_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{SS_{xy}}{SS_{xx}} \quad (\text{EC. 50})$$

$$\beta_0 = \bar{Y} - \beta_1 \bar{X} = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (\text{EC. 51})$$

Para simplificar la notación podemos usar:

La suma de cuadrados de las desviaciones para:

$$SS_{xx}^2 = \sum (x_i - \bar{x})^2 \quad (\text{EC. 52})$$

$$SS_{yy}^2 = \sum (y_i - \bar{y})^2 \quad (\text{EC. 53})$$

Suma de los productos de las desviaciones:

$$SS_{xy}^2 = \sum (x_i - \bar{x})(y_i - \bar{y})$$

Varianza de ambas varianzas

$$s_x^2 = \frac{SS_{xx}}{n-1} \quad (\text{EC. 54})$$

$$s_y^2 = \frac{SS_{yy}}{n-1} \quad (\text{EC. 55})$$

Covarianzas

$$s_{xy}^2 = \frac{SS_{xy}}{n-1} \quad (\text{EC. 56})$$

La línea de mínimos cuadrados puede ser representada por

$$Y = \beta_0 + \beta_1 X \quad (\text{EC. 57})$$

o equivalente a

$$Y = Y + \beta_1(X - X) \text{ (EC. 58)}$$

Las asunciones acerca de e_i :

- 1) e_i es una variable aleatoria normalmente distribuida con media $\mu=0$ y varianza σ^2 ($e_i: N(0, \sigma^2)$).
- 2) Los valores de e_i son independientes.
- 3) Todos los e_i tienen la misma varianza σ^2 . Donde σ^2 es constante es todo el rango de concentraciones. Además es independiente de la concentración (homocedasticidad).

2.7.2.2.3 Medida de la Calidad de la Línea Recta y del Estimador de σ^2

Una vez que los parámetros b_0 y b_1 son determinados por mínimos cuadrados, la línea ajustada se usa para predecir a Y, y en que extensión. Una medida que ayuda a responder es pregunta es calculando la suma de cuadrados del error (SCE).

$$SCE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \text{ (EC. 59)}$$

Hay dos factores que contribuyen a que SCE sea grande:

- (1) Una variación grande los datos, esto es un valor de σ^2 grande.
- (2) El modelo lineal asumido es inapropiado.

Si el segundo factor es falso, se puede obtener el valor de σ^2 (de la población o de la verdadera línea recta), es dado por:

$$S_{y/x}^2 = \frac{1}{n-2} SCE = \frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \text{ (EC. 60)}$$

2.7.2.2.4 Análisis de Varianza

Una manera de evaluar el grado de ajuste de la recta es comparando los datos observados (Y) con los puntos del modelo (\hat{y}), frente a la media de todos los puntos (\bar{y}).

Para cada valor observado y_i , es calculado la diferencia entre \bar{y} y y_i o la desviación total que es la distancia vertical entre el y_i y \bar{y} .

$$\text{desviación total} = y_i - \bar{y}$$

la suma de cuadrados totales (SCT) es obtenida:

$$SCT = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (\text{EC. 61})$$

La desviación explicada es la distancia vertical entre la línea media de los puntos, \bar{y} y un punto sobre la línea de regresión, \hat{y}_i .

$$\text{desviación explicada} = \hat{y}_i - \bar{y}$$

la suma de los cuadrados explicados o regresionales (SCR),

$$SCR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (\text{EC. 62})$$

En general tenemos

$$\begin{array}{ccccccc} y_i - \bar{y} & = & \hat{y}_i - \bar{y} & + & y_i - \hat{y}_i \\ \text{Desviación total} & & \text{Desviación} & & \text{Desviación no} \\ & & \text{explicada} & & \text{explicada} \end{array}$$

$$(y_i - \bar{y})^2 = (\hat{y}_i - \bar{y})^2 + (y_i - \hat{y}_i)^2$$

Suma total de Desviación total		Desviación explicada		Desviación no explicada
SCT	=	SCR	+	SCE
Desviación total		Desviación explicada		Desviación no explicada

La suma de cuadrados no explicados o suma de los cuadrados del error (SCE) es una medida de la dispersión de los datos observados y_i respecto a la línea de regresión.

2.7.2.2.5 Coeficiente de Determinación (R^2):

Intuitivamente si la línea describe bastante bien los la relación entre las variables, el SCR debe ser un numero bastante grande en relación a SCE que debe ser pequeño. La relación entre SCR y SCT es una medida del ajuste de los datos observados al modelo lineal calculado, conocido como coeficiente de determinación.

$$R^2 = \frac{SCR}{SCT} = \frac{\sum_{i=1}^n y_i - \bar{y}}{\sum_{i=1}^n y_i - \bar{y}}^2 \quad (\text{EC. 63})$$

El valor más grande de R^2 es 1, este resultado ocurre cuando toda la variación de y_i es explicada por la regresión y todas las observaciones caen en la línea de regresión. Por otro lado, el más bajo valor de R^2 es 0, esta situación ocurre no es posible explicar la variación de y_i por la regresión y la línea dibujada coincide con y .

2.7.2.2.6 Inferencia de la existencia de la pendiente (β_1) en una línea de regresión.

Usando el test estadístico F. Sea la hipótesis

$$H_0: \beta_1 = 0 \quad \text{y} \quad H_A: \beta_1 \neq 0$$

$$\alpha \leq 0.05$$

Tabla 5. Esquema de tabla anova para una regresión lineal simple

Fuente de variación	SC	d.f	MS	F
Regresión Lineal	SCR	1	$MSR = \frac{SCR}{1}$	$F = \frac{MSR}{MSE}$
Residual	SCE	n-2	$MSE = \frac{SCE}{n-2}$	
Total	SCT	n-1		

2.7.2.2.7 Usando el test estadístico T.

Sea la hipótesis

$H_0: b = \beta_1$ y $H_A: b_1 \neq \beta_1$

Para evaluar la hipótesis se puede usar la distribución z o la t. El test estadístico expresa la relación de una variable distribuida aleatoriamente menos su media dividida por un estimado de su error estándar.

$$t_{exp} = \frac{b_1 - \beta_1}{SE_{b_1}} \text{ (EC. 64)}$$

t = distribución t con n-2 grados de libertad de los valores normalizados de b₁

SE_{b₁} = error estándar de b₁ de la muestra

$$z = \frac{b_1 - \beta_1}{\sigma_{b_1}} \text{ (EC. 65)}$$

z = distribución z de los valores normalizados de b₁

σ_{b₁} = error estándar poblacional

$$\sigma_{b_1} = \frac{\sqrt{\frac{\sigma_{y/x}^2}{\sum (x_i - \bar{x})^2}}}{S_x} = \frac{\sigma_{x/y}}{S_x \sqrt{n-1}} \text{ (EC. 66)}$$

En vista que el error estándar de la población σ_{b₁} es desconocido, se puede estimar el error estándar de la muestral S_{b₁}

$$SE_{b_1} = \frac{\sqrt{S_{y/x}}}{x_i - \bar{x}} \quad (\text{EC. 67})$$

$S_{y/x}^2$ es la varianza de la muestra para y_i , y es un estimador de $\sigma_{y/x}^2$, el cual es desconocido.

$$S_{y/x}^2 = \frac{1}{n-2} \sum (y_i - \hat{y}_i)^2 = \frac{1}{n-2} \sum e_i^2 = \frac{1}{n-2} SCE = MSE \quad (\text{EC. 68})$$

2.7.2.2.8 Inferencia de la existencia de la pendiente (β_0) en una línea de regresión.

Sea la hipótesis

$$H_0: b_0 = \beta_0 \quad \text{y} \quad H_A: b_0 \neq \beta_0$$

$$t = \frac{b_0 - \beta_0}{SE_{b_0}} \quad (\text{EC. 69})$$

La variable aleatoria tiene una distribución t con n-2 grados de libertad.

$$Z = \frac{b_0 - \beta_0}{\sigma_{b_0}} \quad (\text{EC. 70})$$

$$SE_{b_0} = \frac{\sqrt{S_{y/x}}}{n} \frac{x_i}{\sqrt{x_i - \bar{x}^2}} = S_{y/x} \frac{\sqrt{\frac{nx^2}{(n-1)S_x^2}}}{n} \quad (\text{EC. 71})$$

S_{β_0} es un estimador del σ_{β_0} poblacional

En general, para una hipótesis de un nivel de significancia α , la hipótesis nula debe ser rechazada si uno de los siguientes eventos ocurre:

$t \geq t_{n-2, 1-\alpha}$ para un test de una cola a la derecha de la curva de distribución $H_A: \beta_1 > \beta_0$ o

$H_A: \beta_0 > \beta_0$

$t \leq t_{n-2, 1-\alpha}$ para un test de una cola a la izquierda de la curva de distribución $H_A: \beta_1 < \beta_0$ o

$H_A: \beta_0 < \beta_0$

$|t| \geq t_{n-2, 1-\alpha/2}$ para un test de dos colas $H_A: \beta_1 \neq \beta_0$ o $H_A: \beta_0 \neq \beta_0$

Donde $t_{n-2,1-\alpha}$ representa el 100(1- α)% de una distribución t con n-2 grados de libertad.

Intervalos de Confianza:

- 1) El intervalo de confianza para el parámetro β_1 esta dado por $b_1 \pm t_{n-2,1-\alpha/2} SE_{b_1}$
- 2) Para el parámetro β_0 el intervalo de confianza esta dado por $b_0 \pm t_{n-2,1-\alpha/2} SE_{b_0}$
- 3) La ecuación de regresión permite predecir un valor de y dado un valor de x, un intervalo de confianza con un porcentaje de predicción del 100(1- α)% se calcula para el valor medio de los valores predichos de y para un valor de x por medio de la siguiente expresión

$$y_0 \pm t_{n-2, 1-\frac{\alpha}{2}} SE_{y_0/x_0} \quad (\text{EC. 72})$$

$$y + b_1 x_0 - x \pm t_{n-2,1-\alpha/2} SE_{y_0/x_0} (\text{EC. 73})$$

$$SE_{y_0/x_0} = S_{y/x} \sqrt{\frac{1}{n} + \frac{(x_0-x)^2}{(n-1)S_x^2}} \quad (\text{EC. 74})$$

Igualmente, el intervalo de predicción obtenido para un valor de y_0 para un valor dado de x_0 para un t con n-2 grados de libertad es calculado

$$y_0 \pm t_{n-2, 1-\frac{\alpha}{2}} SE_{y_0/x_0} \quad (\text{EC. 75})$$

$$SE_{y_0/x_0} = S_{y/x} \sqrt{1 + \frac{1}{n} + \frac{(x_0-x)^2}{(n-1)S_x^2}} (\text{EC. 76})$$

Visualmente hay una diferencia entre la banda del intervalo de confianza para la media de los valores predichos de y_0 dado un valor x_0 respecto al intervalo de confianza de un valor individual y_0 dado un valor de x_0 .

2.7.3 Modelos de Regresión Lineal Múltiple

El método de mínimos cuadrados (MMC) es usado para estimar los parámetros en modelos de regresión múltiple.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_{ik} x_{ik} + \varepsilon_i \text{ (EC. 77)}$$

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon \text{ (EC. 78)}$$

$i = 1, 2, \dots, n$ Donde n = número de observaciones

$j = 1, 2, \dots, k$ Donde k = número de variables

En este modelo se asume que el error ε en el modelo tiene una media $E(\varepsilon)=0$ y una varianza de $V(\varepsilon)=\sigma^2$, por lo tanto ε es una variable aleatoria.

La suma de los errores al cuadrado L es minimizada para la selección de los coeficientes β 's en la ec. 76.

$$L = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij})^2 \text{ (EC. 79)}$$

$$y_{11} = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \dots + \beta_{1k} x_{1k} \text{ (EC. 80)}$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \dots + \beta_{2k} x_{2k}$$

...

...

$$y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \dots + \beta_{nk} x_{nk}$$

El sistema de ecuaciones puede expresarse en forma matricial:

$$\begin{matrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{matrix} = \begin{matrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{matrix} \begin{matrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{matrix} + \begin{matrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{matrix} \text{ (EC. 81)}$$

Donde:

$$y = X\beta + \varepsilon$$

y = vector columna ($n \times 1$) de observaciones

X = matriz ($n \times k$) de variables independientes

β = vector columna ($k \times 1$) de coeficientes de regresión

ε = vector columna (n x 1) de errores aleatorios

Se desea encontrar el vector de estimadores de mínimos cuadrados β que minimice L.

$$L = \sum_{i=1}^n \varepsilon_i^2 = \varepsilon' \varepsilon = (y - X\beta)' (y - X\beta) \quad (\text{EC. 82})$$

El cual puede expresarse como

$$L = y'y - \beta'X'y - y'X\beta + \beta'X'X\beta \quad (\text{EC. 83})$$

$$L = y'y - 2\beta'X'y + \beta'X'X\beta \quad (\text{EC. 84})$$

Debido a que $\beta'X'y$ es una matriz (1x1) o un escalar y su transpuesta $(\beta'X'y)' = y'X\beta$ corresponde al mismo escalar.

Los estimadores de β se determina al minimizar L, al derivar L respecto β y se iguala a cero.

$$\frac{\partial L}{\partial \beta} = -2X'y + 2X'X\beta = 0 \quad (\text{EC. 85})$$

Se simplifica a

$$X'X\beta = X'y \quad (\text{EC. 86})$$

Se multiplica ambos lados de la ecuación por $(X'X)^{-1}$, el inverso de $X'X$. El estimador de mínimos cuadrados de β es determinado

$$\beta = (X'X)^{-1}X'y \quad (\text{EC. 87})$$

$(X'X)^{-1}$ = matriz simétrica (k x k)

$X'y$ = vector columna (k x 1)

El modelo de regresión será

$$y = X\beta \quad (\text{EC. 88})$$

$$y = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon \quad (\text{EC. 89})$$

El valor residual e_i (n x 1) es la diferencia entre la observación original y_i el valor calculado \hat{y}_i

$$e_i = y_i - \hat{y}_i$$

Estimación de σ^2

La suma de cuadrados de los residuales es

$$SS_E = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n e_i^2 = \varepsilon' \varepsilon \quad (\text{EC. 90})$$

Sustituyendo y por $X\beta$ nos queda

$$e = y - \hat{y} = y - X\beta \quad (\text{EC. 91})$$

Ahora tenemos

$$SS_E = (y - X\beta)'(y - X\beta) \quad (\text{EC. 92})$$

$$SS_E = y'y - \beta'X'y - y'X\beta + \beta'X'X\beta \quad (\text{EC. 93})$$

$$SS_E = y'y - 2\beta'X'y + \beta'X'X\beta \quad (\text{EC. 94})$$

Debido a que $X'X\beta = X'y$ esta última ecuación se convierte en

$$SS_E = y'y - 2\beta'X'y + \beta'X'X\beta \quad (\text{EC. 95})$$

$$SS_E = y'y - \beta'X'X\beta \quad (\text{EC. 96})$$

La suma de cuadrados de los residuales (SS_E) tiene asociado $n - k$ grados de libertad. El estimador de la SS_E es mostrado a continuación

$$E(SS_E) = \sigma^2(n-k)$$

Así que el estimador de σ^2 está dado por

$$\sigma^2 = \frac{SS_E}{n-k} \quad (\text{EC. 97})$$

Tabla 6. Esquema general de tabla anova de una regresión lineal simple

Fuente de variación	SC	d.f	MS	F
Regresión lineal	SCR	K	$MSR = \frac{SCR}{k}$	$\frac{MSR}{MSE}$
Residual	SCE	n-k-1	$MCE = \frac{SCE}{n-k-1}$	
Total	SCT	n-1		

2.7.3.1 Regresión por componentes principales

La regresión por componentes principales (PCR) es una extensión del análisis de PCA. PCR tiene dos pasos, el primer paso es realizar un PCA en la matriz de datos (X). Luego de aplicar el PCA a la matriz X , son obtenidos los scores T que contienen el resumen de la variación en las muestras. La regresión lineal múltiple es lograda usando los scores del PCA como variables de predicción y una matriz Y que contiene la variable de interés para cada muestra. La predicción de la propiedad Y puede ser descrita con la ecuación como sigue:

$$Y_l = b_0 + \sum_{k=1}^i b_k T_k + e_{kl} \quad (\text{EC. 98})$$

Donde b_k son los coeficientes computados, T_k son los scores del análisis de componente principal y e_{kl} es el error de la regresión lineal múltiple.

2.7.3.2 Regresión por mínimos cuadrados parciales

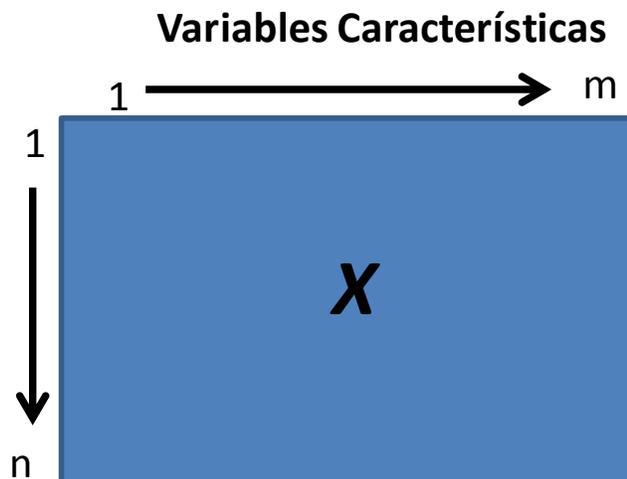
Mínimos cuadrados parciales (PLS: Partial Least Squares) es un método de análisis multivariado usado con frecuencia en calibración multivariado para la determinación de la concentración de un analito para diversas matrices por técnicas como espectrometría de infrarrojo cercano (NIR), espectroscopia de resonancia nuclear, en la cual métodos de regresión lineal o múltiple no son apropiado aplicar debido a la alta colinearidad (correlación muy alta entre las variables) entre las variables descriptivas disponibles para construir el modelo de calibración.

Los PLS son importante como método de calibración porque permiten reducir el número de variables, analizar patrones o para extraer a información útil de grandes fuentes de datos puesta en forma de matrices construidas con los espectros resultante del análisis instrumental. El PLS fue desarrollado por Hernan Wood en 1975 para aplicaciones

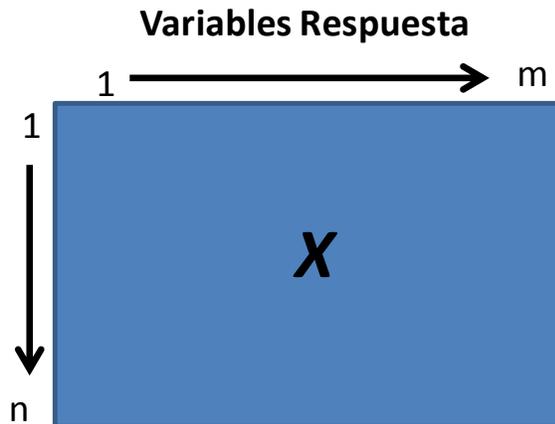
específicas en econometría. Luego su hijo, Steven Wood y otros lo han aplicado exitosamente en la química (Geladi y Kowalski, 1986). Los PLS son aplicados en muchos problemas que involucran la medición de la concentración de monómeros en copolímeros, propiedades fisicoquímica de polímeros que incluyen densidad y fuerza Tensile; índice de octano y contenido de etanol/metano en gasolinas; determinación del contenido de proteína, humedad y grasa en productos agrícolas o alimentos; así como la determinación de la concentración de principios activos y propiedades fisicoquímicas en formas farmacéuticas. Los PLS buscan el hiperplano de máxima varianza entre las variables de predicción (X) y las variables de respuesta (Y), por lo tanto, a diferencia de los PCA, los PLS son conocidos como modelo de factores bi-lineares. Los PLS son usado fundamentalmente para encontrar una relación entre las matrices de variables X y Y.

2.7.3.3. Datos para el Análisis PLS

Sea $X_{n,m} = [x_1, x_2, \dots, x_i, \dots, x_m]$ = una matriz con n filas (objetos) y m columnas (variables descriptivas o de predicción, x_i). Las variables descriptivas son consideradas como vectores columnas y el vector fila es considerado como la traspuesta del vector columna.



Sea $Y_{n,k} = [y_1, y_2, \dots, y_i, \dots, y_k]$ una matriz con n filas (objetos) y k columnas (variables de respuestas, y_i)



2.7.3.4 Pre-procesamiento o Escalamiento de las Variables

En vista de que las variables descriptivas son medidas en diferentes unidades se hace necesario unificar el bloque de variables a una misma escala. Los métodos de escalamiento más usados son:

- a) **Media Centrada.** Los valores de las variables características en la matriz X son centrado, haciendo la media de cada columna igual a cero.

$$x_{ij}(\text{centrado}) = x_{ij} \text{ original} - x_j$$

- b) **Autoescala (transformación-z):** Los valores de x se escalan para que todas las columnas tengan una media de cero y una desviación estándar de 1.

$$x_{ij}(\text{autoescala}) = \frac{x_{ij} \text{ original} - x_j}{s_j} \quad (\text{EC. 99})$$

2.7.3.5 Algoritmo de PLS

El objetivo del PLS es proyectar la información de los datos en un reducido número de nuevas variables llamadas scores de X denotado por t_i y scores de Y denotados por u_i para lo cual $i = 1, 2, \dots, a$, donde a es el número de variables latentes usadas para desarrollar

un modelo de regresión lineal entre t_i y u_i . Los escores de X es un numero pequeño a que son ortogonales entre sí. La combinación lineal entre los escores de X y Y como sigue:

$$u_1 = t_1^T \cdot b_1 + e_1 = t_{11}b_{11} + t_{12}b_{21} + \dots + t_{1i}b_{i1} + \dots + t_{1a}b_{p1} + e_1 \quad (\text{EC. 100})$$

..

$$u_j = t_j^T \cdot b_j + e_j = t_{j1}b_{12} + t_{j2}b_{22} + \dots + t_{ji}b_{i2} + \dots + t_{ja}b_{a2} + e_j \quad (\text{EC. 101})$$

..

$$u_n = t_n^T \cdot b_n + e_n = t_{n1}b_{1p} + t_{n2}b_{2p} + \dots + t_{ni}b_{ip} + \dots + t_{na}b_{an} + e_n \quad (\text{EC. 102})$$

Escrito en forma matricial

$$U = TB + E \quad (\text{EC. 103})$$

Donde E es un vector de errores y B es un parámetro desconocido que se puede estimar mediante la siguiente expresión de regresión

$$B = T^T T^{-1} \cdot T^T U \quad (\text{EC. 104})$$

La matriz $T_{n,m}$ son los escores de X y la matriz $U_{n,m}$ son los escores de Y . La matriz de escores T y U contienen información de la regresión lineal entre la matriz X y la matriz Y .

Los escores de X (t_{ia}) y Y (u_{ia}) son calculados como sigue

$$T = XW_x = t_{ia} = \sum_{i=1}^n Xw_j \quad (\text{EC. 105})$$

Donde la matriz de transformación de pesos W_x puede calcularse así

$$W_x = W(P^T W)^{-1} \quad (\text{EC. 106})$$

Donde W es la matriz de pesos de X .

y

$$U = YC = u_{ia} = \sum_{j=1}^n Yc_j \quad (\text{EC. 107})$$

Donde w_j y c_j son vectores con una longitud igual a la unidad determinados maximizando la covarianza entre t_j y u_j así

$$w = \frac{\overline{w^T w}}{w^T w} = 1 \quad (\text{EC. 108})$$

Y

$$c = \frac{\overline{c^T c}}{c^T c} = 1 \quad (\text{EC. 109})$$

En general los escores de X se debe cumplir con las siguientes propiedades

- 1) La variables de predicción $X = [x_1, x_2, \dots, x_i, \dots, x_m]$ puede ser predicha bastante al multiplicar T (los escores de X) por P que corresponde a los loading, así que los residuales de X, e_{ik} son pequeños.
- 2) $X = TP^T + E$ o lo que es equivalente

$$x_{im} = \sum_{i=1}^m t_{ia} \cdot p_{ak} + e_{ik} \quad (\text{EC. 110})$$

Donde T es la matriz de escores de X ; P es la matriz de loading y E es el vector de residuales de X.

- 3) La matriz de predicción $Y = [y_1, y_2, \dots, y_i, \dots, y_m]$ es calculada al multiplicar U (escores de Y) por la matriz de pesos o ponderación C, para la cual G es la matriz de residuales de Y.
- 4) $Y = UC^T + G$ o su escrito en forma equivalente

$$y_{im} = \sum_{i=1}^m u_{ia} \cdot c_{am} + g_{im} \quad (\text{EC. 111})$$

Donde U es la matriz de escores de Y; C es la matriz de pesos y G es el vector de residuales de Y.

Igualmente, debido a que los escores de X son buenos predictores de Y , también pueden predecir bastante bien a la matriz Y en términos de los escores de X con su correspondiente matriz de residuales F .

$$Y = TC^T + F \quad (\text{EC. 112})$$

Equivalente a

$$y_{im} = \sum_{i=1}^m t_{ia} \cdot c_{am} + c_{im} \quad (\text{EC. 113})$$

Los residuales E , F y G expresan la desviación entre los valores observados y la respuesta de los valores modelados.

Las ecuaciones pueden describirse como un modelo de regresión lineal múltiple para predecir la matriz de respuesta Y así:

$$Y = TC^T + F = XW C^T + F = XWC^T + F \quad (\text{EC. 114})$$

$$Y = XB + F \quad (\text{EC. 115})$$

$$B = WC^T \quad (\text{EC. 116})$$

2.7.3.6 Algoritmo NIPALS

El algoritmo NIPALS (Non-Linear Iterative Partial Least Squares) más simple para calcular los parámetros del PLS desarrollado por Wood y col. El algoritmo NIPALS trabaja con las matrices originales X y Y escalada o centradas.

1. Introduce las matrices X y Y
2. se asigna un vector columna w y a u

$$w = y_1$$

3. Se calcula w , los pesos de X

$$w = X^T u \frac{1}{u^T u} \text{ (EC. 117)}$$

4. Calcular t , los escores de X

$$t = X^T w \frac{1}{w^T w} \text{ (EC. 118)}$$

5. Calcular c , los pesos de Y

$$c = Y^T t \frac{1}{t^T t} \text{ (EC. 119)}$$

6. Calcular el nuevo u actualizando sus valores

$$u = Y^T c \frac{1}{c^T c} \text{ (EC. 120)}$$

7. probar si hay convergencia para un $\epsilon \leq 10^{-6}$ o 10^{-8}

Si $\frac{t_{i+1} - t_i}{t_{i+1}} \gg \epsilon$ **Entonces**

Vaya al paso 3

Si No calcular

$$p = X^T t \frac{1}{t^T t} \text{ (EC. 121)}$$

$$X = X - t.p^T \text{ (EC. 122)}$$

$$Y = Y - t.c^T \text{ (EC. 123)}$$

Termina Si

8. **Si** variable Y es solo una columna, $m=1$ **Entonces**

Vaya al paso 9

Si No

Vaya al paso 2

Termina Si

9. **Termine programa**

2.7.3.7 Validación del Modelo y Selección del Número de Componentes

El paso final en un modelo de PLS es la determinación del número de componentes o factores (A) a utilizar en la predicción de las propiedades Y, además el número de componentes A en el modelo es una indicación de la complejidad de este. Esto implica que es necesario calcular el número de componentes A que correlacionen con la variable X y no se produzca un sobre estimado ajuste (over-fitting) que no tiene un verdadero poder de predicción del modelo. Por lo tanto el número de componentes A que contienen información relevante es determinado, así mismo es validado antes de usarse para predecir ya sea propiedades biológicas o fisicoquímicas. El método por excelencia usado para validar y calcular los componentes óptimos de un modelo de PLS es la validación cruzada (CV, cross validation).

2.7.3.7.1 Validación Cruzada (CV)

La CV divide todo el conjunto de datos en 5 a 9 grupos escogidos al azar y desarrollando modelos paralelos donde es descartado uno de los grupos cada vez. Una vez es corrido cada modelo es calculado la diferencia entre los valores de Y, el valor actual (Y_i) y el predicho ($Y_{i,pred}$). La suma del cuadrado de estas diferencias es calculada y obtenida para todos los modelos paralelos hasta formar la suma residual predictiva de cuadrados (PRESS, Predictive Residual Sum of Squares) en la cual este valor estima la habilidad de predicción del modelo.

$$PRESS = \sum_{i=1}^n (y_i - y_{i,pred})^2 \text{ (EC. 124)}$$

Una medida de la predicción de la bondad de la predicción puede también realizarse mediante validación cruzada de R^2 expresado por Q^2 , la cual es calculada así

$$Q^2 = 1 - \frac{PRESS}{SS} \text{(EC. 125)}$$

$$SS = \sum_{i=1}^n (y_i - y_{i,pred})^2 \text{(EC. 126)}$$

Donde $Y_{i,pred}$ en la ecuación anterior es estrictamente calculado como el valor medio de los valores del grupo de validación cruzada más que el valor medio de todo el conjunto de datos.

2.7.3.7.2 Coeficiente de Determinación (R^2)

El coeficiente de correlación, R^2 es un valor está entre 0 y 1 e indica la proporción de la variación en la variable independiente que es explicada por la ecuación de regresión. El R^2 es calculado mediante la siguiente expresión:

$$R^2 = \frac{SCE}{SCT} = \frac{SCT - SCR}{SCT} = 1 - \frac{SCR}{SCT} \text{(EC. 127)}$$

$$SCT = SCR + SCE \text{(EC. 128)}$$

Suma de Cuadrados Totales:

$$(SCT) = \sum (y_i - \bar{y})^2 \text{(EC. 129)}$$

Suma de Cuadrados Explicada:

$$(SCE) = \sum (\hat{y}_i - \bar{y})^2 \text{(EC. 130)}$$

Suma de Cuadrados Residuales:

$$(SCR) = \sum (y_i - \hat{y}_i)^2 \text{(EC. 131)}$$

Un R^2 corresponde a un valor de cero en la situación en la cual ninguna de la variación en las observaciones de la variable dependiente es posible explicar mientras un valor de uno cuando corresponde a una explicación completa. R^2 es una medida del grado de ajuste de

los datos al modelo de regresión por PLS mientras que Q^2 es determina la habilidad de predicción es reflejada en el valor de R^2 . El valor de R^2 es calculado usando la totalidad de los datos mientras que para calcular Q^2 los datos son divididos en 4 o más grupos, en la cual siempre se deja a fuera uno de estos grupos. Si hay una gran discrepancia entre R^2 y Q^2 , en la cual el valor R^2 es significativamente mayor que el valor de Q^2 (25%), es una indicación que hay una sobre estimación del ajuste (over-fitting) de los datos por el modelo y la habilidad de predicción no es adecuada.

2.7.3.7.3 Error Cuadrado de la Raíz Media (RMSE, Root-Mean Squared Error)

El RMSE al igual que PRESS y Q^2 estima la capacidad predictiva del modelo para establecer si R^2 es sobre estimado (over-fitting). El RMSE muestra el error entre el valor de la media experimental y la de predicción.

$$RMSE = \frac{\sqrt{\sum_{i=1}^n (y_{i,pred} - y_i)^2}}{n} \text{ (EC. 132)}$$

Donde $y_{i,pred}$ es la predicción de la variable dependiente y n es el número de objetos o moléculas en el conjunto de datos. Valores de RMSE mayores o iguales a 1 refleja una pobre habilidad del modelo de predecir con exactitud los datos a pesar que se reporte un valor de R^2 cercano a uno (≥ 0.75).

2.7.3.8 Evaluación de los modelos de predicción

Durante el desarrollo de modelos de calibración, el número de factores que el modelo requiere debe ser determinado. Uno de los criterios establecidos para la elección del número de factores es siempre obtener el menor error de predicción. El error cuadrático medio de predicción (RMSEP) es utilizado para evaluar los errores de calibración y los modelos de predicción, la ecuación 134, indica como calcularlo.

$$\text{RMSEP} = \frac{\sum_{i=1}^n (Y_{\text{predicho}} - Y_{\text{aceptado}})^2}{n} \quad (\text{EC. 133})$$

Donde n es el número de muestras, Y_{predicho} el valor de la propiedad predicho por el modelo, Y_{aceptado} el valor de la propiedad aceptado como verdadero (Reich, 2005).

2.7.3.8.1 Estrategias Generales en el Análisis

A continuación se describe la estrategia general que suele seguirse en la construcción de los modelos, que si bien es similar tanto para el análisis cualitativo como para el cuantitativo, cada uno de ellos tiene algunas particularidades. Teniendo en cuenta las características del sistema objeto de estudio y lo que se pretende determinar, se selecciona un modelo quimiométrico (métodos cualitativos y/o cuantitativos) (Reich, 2005, Bakeev, 2005). Siguiendo los pasos enunciados a continuación:

1. Examinar los datos. Hacer un análisis de los datos para ver si existen errores burdos. Estos errores pueden producirse tanto en las variables medidas (Ej. Curvas FID) como en los valores de referencia (propiedades físicas y/o químicas), por lo que es importante examinar ambas variables.
2. Selección de las muestras de los conjuntos de calibración y predicción. Este es un aspecto de especial relevancia, ya que el conjunto de calibración debe estar constituido por muestras que cubran el rango del parámetro a determinar y también la variabilidad espectral. Una forma habitual de hacer esta selección es a partir del gráfico de los scores del PCA. También se ha de seleccionar un segundo conjunto de muestras que conformaran el conjunto de predicción, con las cuales se evaluará la capacidad predictiva del modelo construido (Bakeev, 2005).
3. Selectividad, la selectividad es la capacidad de un método para analizar el analito de interés y no otro componente. Los modelos de calibración se desarrollan con

muestras que contienen el analito a cuantificar y el resto de componentes de la muestra. Previa a la cuantificación se realiza un análisis cualitativo, que consiste en una identificación mediante un análisis de PCA preparado con muestras que pertenecen a la misma clase. Con esto se garantiza que las muestras que van a ser cuantificadas, pertenecen al grupo de interés (Bakeev, 2005).

4. Linealidad, es demostrada obteniendo una respuesta proporcional entre la señal y la propiedad a determinar, en todo el intervalo. En las calibraciones univariadas la linealidad se obtiene una relación lineal entre la señal y la concentración. En las calibraciones multivariadas se evalúa la linealidad mediante la relación entre el valor de referencia y el valor estimado con el método desarrollado. Se recomienda tener un mínimo de 5 niveles de concentración que cubran todo el rango (Bro, 2003).
5. Exactitud, la exactitud se evalúa por la semejanza entre el valor estimado y el valor de referencia o aceptado, ya sea este último obtenido por pesada o por un método de referencia. La exactitud quedara demostrada con un test-t de las diferencias, donde el promedio de estas diferencias deber ser significativamente igual a cero a un nivel de significación establecido.
6. Precisión, la precisión es el error aleatorio o indeterminado asociado al método de análisis. Puede ser expresada como la desviación estándar relativa o coeficiente de variación (%).

3. MATERIALES Y MÉTODOS

3.1 MUESTRAS EN FORMA DE POLVO Y DE PELLETS

En este estudio se emplearon muestras de PP-h (polímero de propileno) y CPE-i, CPE-LER, CPE-r (polímero Propileno-etileno). Todas las muestras fueron obtenidas en una planta comercial de producción de polipropileno (PP).

Las muestras de PP-h, CPE-LER y CPE-r fueron producidas a través de un proceso de polimerización con un solo reactor de lecho fluidizado. Los copolímeros de impacto fueron producidos en dos reactores de lecho fluidizado ubicados en serie; un primer reactor produjo la resina homopolimérica que luego fue alimentada a un segundo reactor para producir el copolímeros de impacto al alimentarse los monómeros etileno-propileno en una relación cercana a 1:1. El sistema catalizador utilizado en la producción de las muestras de PP fue del tipo Ziegler-Natta.

Las muestras de los diferentes PP en polvo seleccionadas poseían un índice de fluidez alrededor de 3.0 g/min (230 °C y 2160 g). Mientras que las muestras de pellets para el desarrollo de los métodos de predicción de propiedades mecánicas tenían un rango más amplio de índices de fluidez.

Las muestras de CPE-LER (n = 200) con un rango de etileno entre 0.1 y 1.1% fueron seleccionadas para generar un modelo de cuantificación del porcentaje de etileno usando resonancia magnética nuclear protónica (¹H-NMR). Las muestras seleccionadas para generar y evaluar el modelo de predicción PLS para la propiedad % Et, fueron muestras de CPE-LER con una concentración de etileno <1.1%. Los análisis al CPE-LER se realizaron sobre las muestras en forma de polvo a la salida del reactor previo al proceso de pelletizado.

Un grupo de 150 muestras de PP-h, CPE-LER, CPE-r y CPE-i, fueron seleccionadas para desarrollar un modelo de predicción de resistencia a la tensión y uno de módulo de flexión usando $^1\text{H-RMN}$ por análisis multivariado el mismo grupo de muestras se usó para desarrollar ambos modelos. Para esto, una porción de cada una de las diferentes muestras de PP en forma de pellets fueron convertidas a especímenes de ensayo tipo I mediante el proceso de moldeo por inyección siguiendo lo establecido en la norma ASTM D-4101, posteriormente acondicionados, y finalmente sometidos a los ensayos de tracción y flexión siguiendo las metodologías descritas en las normas ASTM D-638 y D-790 respectivamente. Las muestras seleccionadas fueron empleadas en la obtención de un modelo de predicción PLS para la determinación de la resistencia a la tensión (Tensile) y del módulo de flexión 1% secante.

3.2 ADQUISICIÓN DE LAS CURVAS DE FID

Las curvas de FID de Resonancia magnética nuclear protónica ($^1\text{H-RMN}$) fueron adquiridos para muestras de polipropileno en forma de pellets y en polvo. Las curvas de FID de $^1\text{H-RMN}$ de las muestras estudiadas en pellets y en polvo, fueron adquiridos en un equipo Magstation II RMN espectrómetro de 20 MHz (Progression, Inc. Haverhill, MA, EE.UU.). El espectrómetro Magstation II $^1\text{H-RMN}$ de baja resolución, opera con el software Az/Tec el cual genera los espectros para cada muestra registrada. Las muestras poliméricas en forma de pellet o polvo, son depositadas en un cilindro con diámetro de 20 mm para la lectura espectral de $^1\text{H-RMN}$. Todas las muestras fueron analizadas a una temperatura de $\approx 80^\circ\text{C}$. Los espectros $^1\text{H-RMN}$ generaron seis índices o descriptores primarios para cada muestra.

3.3 DETERMINACIÓN DEL PORCENTAJE DE ETILENO USANDO INFRARROJO MEDIO

Espectros de infrarrojo medio (MIR) se obtuvieron para determinar el porcentaje de etileno (%Et) en 200 muestras de CPE-LER. Los espectros de MIR fueron adquiridos a partir de películas o láminas delgadas de 10 milésimas de pulgada de espesor fabricadas mediante

prensado a 230°C de cada muestra de resina en polvo de CPE-LER. Los espectros MIR fueron adquiridos utilizando un espectrómetro Infrarrojo con transformada de Fourier (FT-IR) Nicolet 6700 (Thermo Electron, EE.UU.) con un rango espectral de 4800-400 cm^{-1} . Todos los espectros se coleccionaron a temperatura ambiente ($\approx 28^\circ\text{C}$), con una acumulación de 64 barridos (Scans), con resolución espectral de 2 cm^{-1} y una velocidad de espejo de 0.6320 cm/s . Los valores de contenido de etileno se calcularon usando una curva de calibración basada en la absorbancia a 732 cm^{-1} correspondiente a la secuencia lineal de trimetileno $-(\text{CH}_2)_3-$ generado por la presencia del etileno en la cadena polimérica.

3.4 MEDICIÓN DE LA RESISTENCIA A TENSIÓN (TENSILE)

Especímenes normalizados para el ensayo de tracción se prepararon mediante moldeo por inyección a partir de 150 muestras de polipropileno en forma de pellets bajo las condiciones y lineamientos establecidos en la norma ASTM D-4101. Fueron especímenes normalizados tipo I con las características y dimensiones descritas en la norma ASTM D-638. Los especímenes-muestra fueron acondicionados por 40 horas a temperatura ambiente de $23 \pm 2^\circ\text{C}$ y una humedad relativa de $50 \pm 5\%$. Una máquina universal de ensayos, Instron modelo 3345 (Norwood, MA, EE.UU.) se empleó en la medición de la resistencia a la tracción (Tensile) y el módulo de flexión 1% secante. Al menos cinco ejemplares por cada muestra fueron preparados y a cada espécimen se le midió el espesor que fue incluido en la fórmula de cálculo empleada por software para determinar el Tensile.

3.4 MEDICIÓN DEL MÓDULO DE FLEXIÓN 1 % SECANTE

Especímenes normalizados de 150 muestras en forma de pellets utilizadas para el ensayo de módulo de flexión 1% Secante fueron preparados mediante moldeo por inyección en las condiciones y lineamientos establecidos por la norma ASTM D-4101. El Módulo de Flexión 1% Secante fue determinado según lo establecido en la norma ASTM D-790 procedimiento A, usando la máquina universal de ensayos Instron 3345 (Norwood, MA, EE.UU.). Se analizaron tres probetas o especímenes de ensayo normalizado por cada

muestra, usando una velocidad de aplicación de la carga de 1.3 mm/min y reportando el promedio de las tres mediciones.

3.5 PRE-PROCESAMIENTO DE LOS DATOS

Las curvas de FID del ^1H -RMN adquiridos para las muestras fueron procesados de dos maneras diferentes. En primer lugar, las variables descriptivas o de predicción fueron obtenidas a partir de la curva FID del ^1H -RMN, la cual es una curva compuesta de la suma de las funciones de decaimiento rápido, decaimiento lento y exponencial. Por otro lado, el método usado para procesar las curvas de FID obtenidas del ^1H -RMN, se realizó usando los coeficientes de las ecuaciones matemáticas usadas para simular la curva de FID con los seis coeficientes mencionados anteriormente: A, a, B, b, C y c. Posteriormente mediante transformaciones y combinaciones de estos coeficientes fueron generados nuevas variables descriptivas, en total fueron obtenidas 230. Las 230 variables descriptivas generadas corresponden a cocientes y productos de términos dobles y triples para los seis coeficientes iniciales de la ecuación matemática del modelo.

Las curvas FID del ^1H -RMN y los descriptores matemáticos adquiridos se estudiaron mediante análisis exploratorio, usando el análisis de componentes principales (PCA). La generación de modelos de predicción para %Et, resistencia a la tensión y módulo de flexión 1%secante, se realizaron usando regresión lineal por mínimos parciales PLS. Para el análisis Multivariado de la información espectral se empleó el software Unscrambler X (versión 10.0, CAMO Software AS, Norway).

3.6 GENERACIÓN Y EVALUACIÓN DE LOS MODELOS DE PREDICCIÓN OBTENIDOS

Los modelos de predicción PLS fueron generados a partir de un grupo de muestras de calibración cuyas propiedades se determinaron usando las mediciones tradicionales tal como se describió en los apartados anteriores. Para el grupo de calibración fueron

seleccionados de manera aleatoria 60% de las muestras; el restante 40 % de las muestras fueron usadas para validación de los modelos creados. Los modelos de predicción obtenidos fueron evaluados mediante el error cuadrático de predicción para muestras de calibración, usando como método de validación interna la validación cruzada y el error cuadrático medio para muestras de validación durante la predicción del conjunto externo de muestras (RMSEC y RMSEP).

4. RESULTADOS

4.1 ESPECTROS ^1H -RMN OBTENIDOS PARA MUESTRAS DE PP HOMOPOLÍMERO Y COPOLÍMERO

Curvas FID/ ^1H -RMN fueron generadas para un grupo representativo de muestras de PP-h, CPE-LER, CPE-r y CPE-i indicando que hay una gran diferencia entre estas en las tres regiones de las curvas: Gaussiana rápida, exponencial y Gaussiana lenta (Figura 14). Las curvas FID/ ^1H -RMN adquiridas en la misma muestra, seleccionada directamente del reactor de producción del sólido de CPE-LER en forma de un polvo y convertida en pellets también fueron comparadas, mostrando solo diferencias en la región Gaussiana rápida (Figura 14(b)). Un análisis comparativo de la forma de las curvas FID/ ^1H -RMN obtenidas para muestras en polvo obtenidas del reactor y muestras de este mismo material en forma de pellets se presenta en la figura 14(b). Estas diferencias son atribuidas a la presencia de cristales con conformaciones desiguales en estas dos presentaciones de PP. De acuerdo a lo anterior, cada grupo de PP puede ser identificado según la inclinación y según la amplitud de la curva FID/ ^1H -RMN.

Tal como se puede apreciar en las curvas de la figura 14, es importante resaltar que a diferencia de otras técnicas espectroscópicas como UV/VIS, Raman o IR medio, las curvas FID/ ^1H -RMN no presentan bandas o picos de absorción discretos, característicos de grupos funcionales. Sin embargo, esto no es una limitación de la capacidad analítica instrumental de esta técnica, tal como se ha mostrado en el apartado anterior, las diferencias son mostradas en la inclinación y por consiguiente en la intensidad de la señal en cada una de las tres regiones de las curvas FID/ ^1H -RMN.

Las curvas FID/ ^1H -RMN de homopolímero y los tres tipos de copolímeros mencionados poseen diferencias muy sutiles en los tres segmentos de la curva de decaimiento (figura 14(a)). Las pequeñas diferencias entre los espectros ^1H -RMN de los tipos de PP, se

atribuyen a su composición química; las muestras de PP-h, CPE-LER, CPE-r y CPE-i, se diferencian por el porcentaje de etileno agregado.

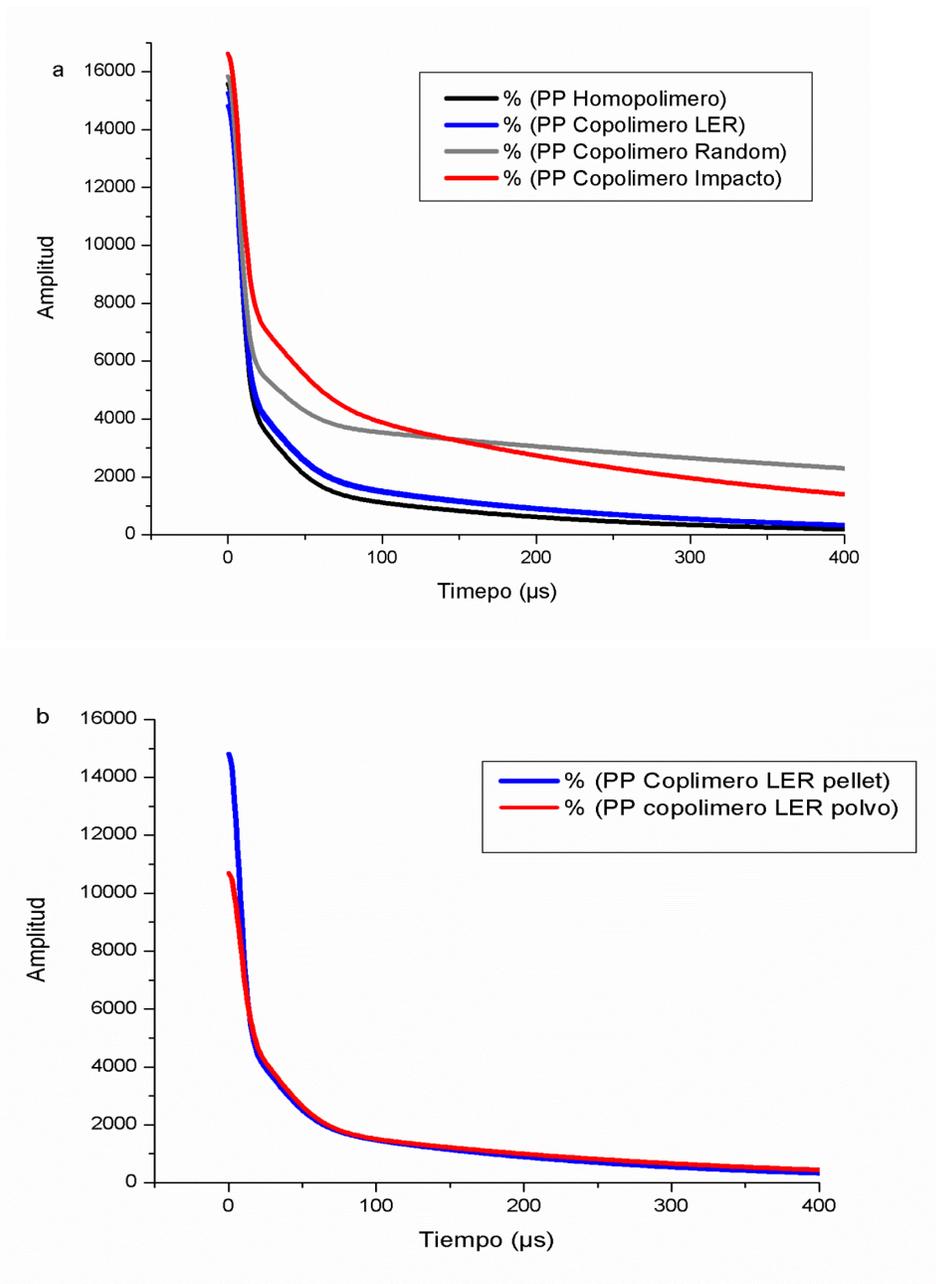


Figura 14. (a) Curvas FID/¹H-RMN generadas para muestras de PP-h, y los diferentes tipos de copolímeros propileno/etileno en forma de pellets (b) Curvas FID/¹H-RMN generadas para muestras de CPE-LER comparando la presentación en forma de polvo y en forma de pellets.

El PP-h no contiene etileno en su estructura polimérica. El porcentaje de etileno agregado al copolímero LER varía entre 0.1 al 1.1%; en cambio el porcentaje de etileno en los PP copolímero random varía entre el 1.2 y el 5.5% mientras que en el PP copolímero de impacto puede llegar hasta un 35% y este a su vez presenta secuencias largas de moléculas de etileno dentro de la cadena polimérica, afectando al mismo tiempo la conformación de los cristales de cada polímero.

Las curvas FID/¹H-RMN en la región Gaussiana de decaimiento rápido para las muestras de CEP-r y CEP-i son las menos pronunciadas. En general, los resultados obtenidos muestran en el análisis de las curvas FID/¹H-RMN diferencias en cada región de acuerdo al tipo de polímero analizado. La primera región de la curva FID (Gaussiana de decaimiento rápido) está asociada con la cristalinidad del polímero, por lo tanto las muestras de PP-h y CEP-LER presentan el mayor alargamiento y por tanto la mayor cristalinidad por poseer la mayor estereoregularidad de las cadenas isotácticas lo cual favorece el agrupamiento de las cadenas en estructuras cristalinas. En los CEP-r y CEP-i se observa menor intensidad en esta zona de la curva FID, lo cual es consecuente con su menor cristalinidad debida a la mayor concentración de etileno en la formación de las cadenas poliméricas.

La segunda región o región central en las curvas FID/¹H-RMN, está asociada a las regiones de transición cristalina-amorfa de la muestra, en esta región las curvas FID de CEP-r y CEP-i presentan mayor intensidad asociada con el mayor contenido de etileno en sus estructuras. El resultado de mayores regiones de transición para los copolímero random y de impacto es consistente con lo esperado según la estructura química conocida para las muestras (Tripathi, 2002).

En la región espectral asociada a la proporción de estructura amorfa para las curvas de FID/¹H-RMN corresponden al decaimiento exponencial. Como es de esperarse, la curva de FID/¹H-RMN refleja un mayor decaimiento exponencial a medida que aumenta la región amorfa en el tipo de muestra, como se presenta en el siguiente orden:

CEP-r > CEP-LER > PP-h

Los copolímeros CEP-i y CEP-r muestran valores de absorción mayores en la región asociada a la proporción de estructura amorfa. Es de resaltar que el CEP-i tiene menor proporción de estructura amorfa que los CEP-r. La razón de menor proporción en estructura amorfa para los copolímeros de impacto reside en la fase homopolímero altamente isotáctica que forma el núcleo estructural de estos copolímeros.

Es posible establecer una diferencia entre las muestras en forma de pellets y las muestras en forma de polvo (Figura 14(b)). En las curvas FID para las muestras en forma de pellets y polvo para el copolímero CEP-LER, presentan diferentes intensidades en la región cristalina de las curvas de FID/¹H-RMN. Para las muestras CPE-LER en polvo y en pellets, la mayor región de diferencia aparece en las intensidades de la Gaussiana rápida. La explicación de la diferencia de intensidades entre las muestras en polvo y las muestras en pellets puede ser atribuida a la mayor probabilidad que tienen las moléculas de cristalizar al formarse los pellets, puesto que las muestras en estado fundido en la etapa de enfriamiento experimentan la masiva agrupación de las moléculas de PP en estructuras cristalinas, en tanto que en el polvo sólo tendrían probabilidad de asociarse en estructuras cristalinas aquellas moléculas que se acaban de formar en cada granulo durante la reacción de polimerización a partir de la partícula de catalizador.

Las diferencias entre los eventos que determinan el proceso de cristalización dan origen a una mayor estructuración de las muestras de PP en forma de pellet en comparación con las muestras de PP en polvo. Para los espectros ¹H-RMN de las muestras de CPE-LER en polvo y en pellets, las regiones amorfas y de transición presentan intensidad y forma similar.

4.2 TRANSFORMACIONES QUIMIOMÉTRICAS APLICADAS A LOS ESPECTROS ^1H -RMN DE MUESTRAS DE PP.

Antes de realizar los modelos de calibración para las propiedades de interés, a saber: a) % de etileno, b) resistencia a la tensión a la cedencia y c) módulo de flexión 1% secante, se aplicaron una serie de pre-tratamientos, entre ellos: a) correcciones de línea base, b) SNV (variación normal estandarizada), c) primera y d) segunda derivada a las curvas FID/ ^1H -RMN con la intención de visualizar mejor la información contenida en los espectros ^1H -RMN. Los resultados de las transformaciones quimiométricas aplicadas a los espectros ^1H -RMN son descritos a continuación.

El primer pre-tratamiento quimiométrico aplicado a las curvas FID/ ^1H -RMN, fue la corrección de línea base. La Figura 15 muestra las curvas FID/ ^1H -RMN de muestras de PP en forma de pellets luego de aplicar la corrección de línea base. La transformación por corrección de línea base permite ver diferencias en la inclinación de las curvas FID/ ^1H -RMN. La aplicación de la corrección de línea base permitió ver una clara tendencia en las regiones de transición y decaimiento exponencial para las curvas FID/ ^1H -RMN. Igualmente, la aplicación de la corrección de línea base a las curvas FID/ ^1H -RMN permite una clara separación de las muestras de CEP-i respecto a las otras tres familias de PP estudiadas.

Se eligió este tipo de corrección de línea base porque permite observar los cambios en la inclinación de la curva FID al colocar un punto fijo y ver la variación en el resto de la curva FID.

El segundo pre-tratamiento quimiométrico aplicado a las curvas FID/ ^1H -RMN, fue el SNV. La Figura 16 muestra las curvas FID/ ^1H -RMN de muestras de PP en pellets luego de aplicar SNV. La transformación por SNV permite ver mejor las tendencias de cada familia de PP en las diferentes regiones de la curva FID ^1H -RMN. El pre-tratamiento SNV acentúa las diferencias en las tres regiones espectrales amorfas, de transición y en la región

cristalina. Además corrige las diferencias entre los espectros debido a efectos de difracción producto de los diferentes tamaños de partícula o de pellets.

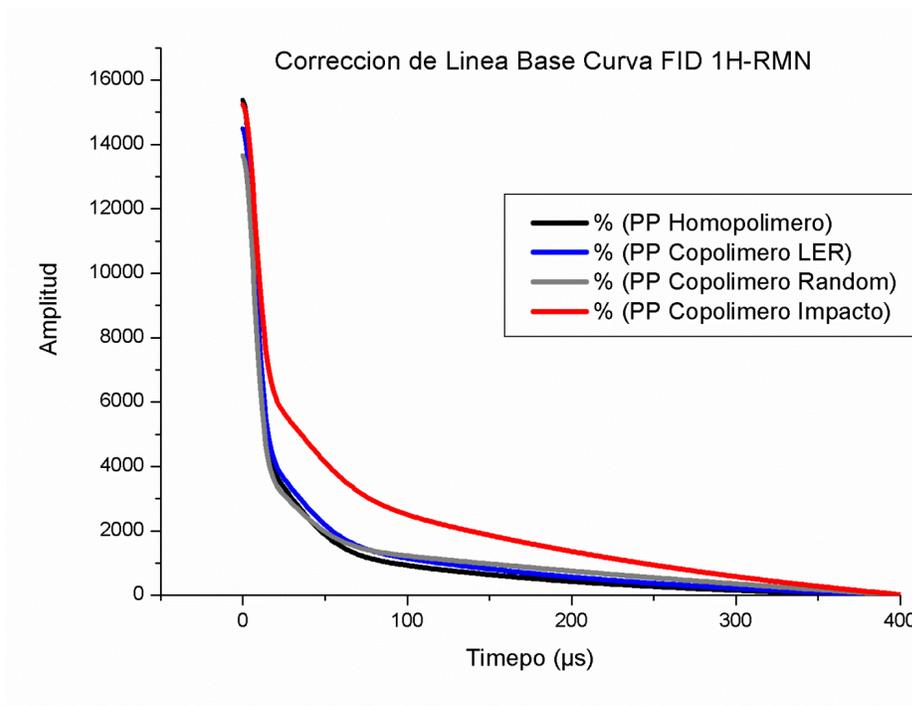


Figura 15. Corrección de línea base aplicada a curvas FID ^1H -RMN de muestras de PP en forma de pellets.

La aplicación de la primera derivada a las curvas FID/ ^1H -RMN, es mostrada en la Figura 15. La transformación por primera derivada permite ver diferencias en la región cristalina de las curvas FID/ ^1H -RMN. La primera derivada, aumentó las diferencias en la región cristalina de los las curvas FID/ ^1H -RMN. El pre tratamiento de la primera derivada disminuye las diferencias en la región amorfa y de transición para las curvas FID/ ^1H -RMN de las muestras estudiadas. Una posible explicación de este resultado es que las curvas FID no presentan una marcada diferencia en la pendiente en la región amorfa de la curva FID.

La aplicación de la segunda derivada a las curvas FID/ ^1H -RMN, crea un efecto de reducción en las diferencias de las curvas FID (Figura 17). Las curvas FID muestran una

apreciable diferencia en la región cristalina y en la región de transición, donde cada una de las regiones las coloca en las coordenadas negativo y positiva del plano cartesiano, respectivamente.

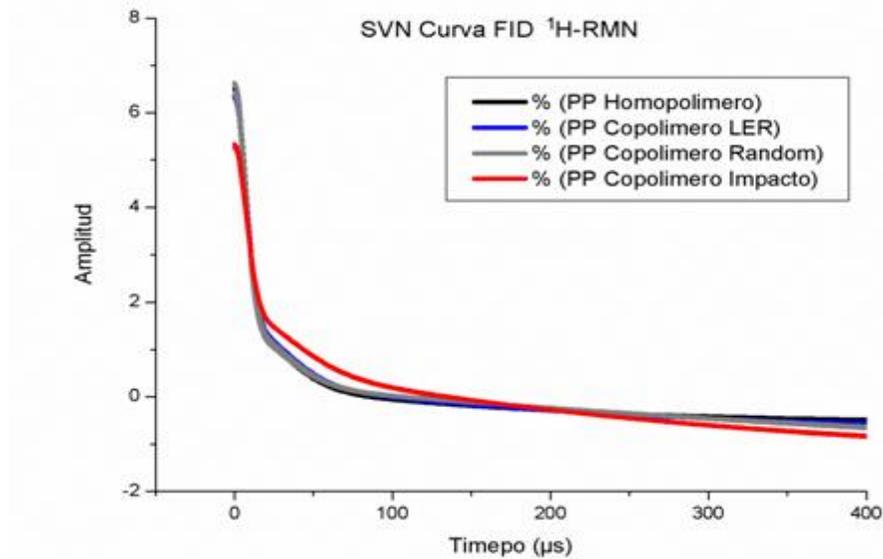


Figura 16. SNV aplicada a curvas FID ¹H-RMN de muestras de PP sólido.

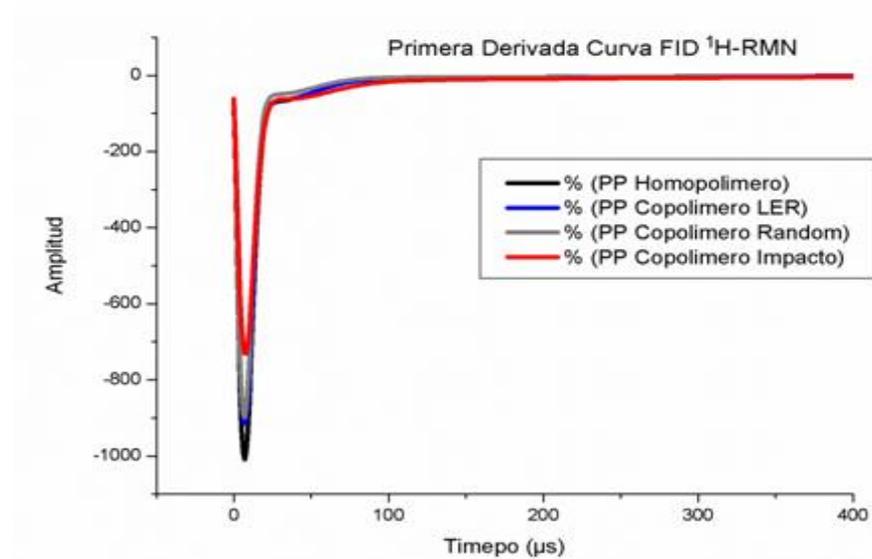


Figura 17. Primera derivada aplicada a curvas FID ¹H-RMN de muestras de PP sólido.

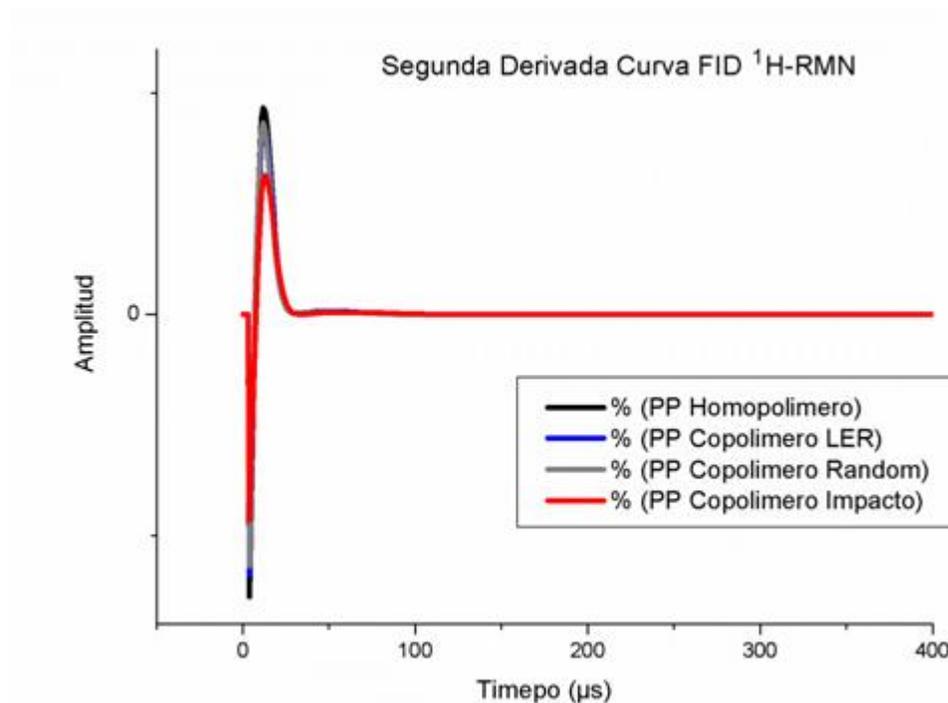
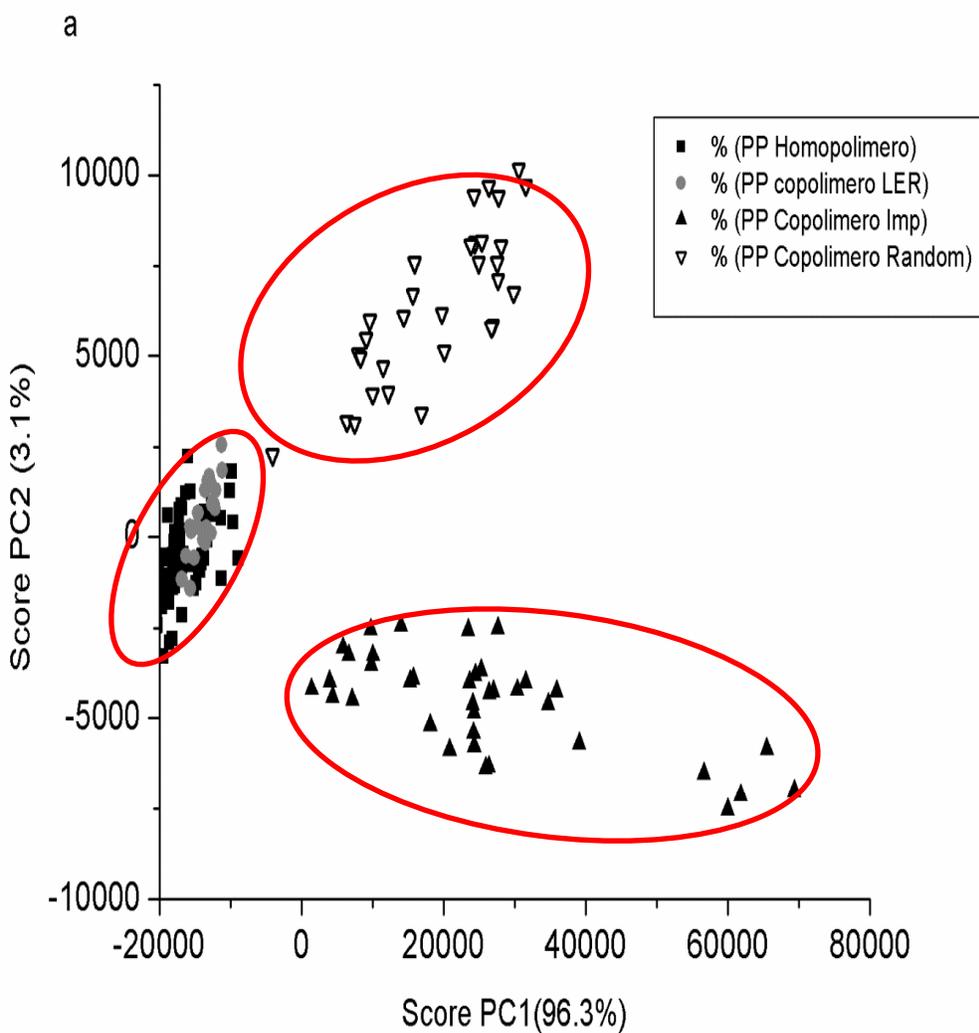


Figura 18. Segunda derivada aplicada a curvas FID $^1\text{H-RMN}$ de muestras de PP sólido

4.3 ANÁLISIS DE COMPONENTES PRINCIPALES DE LAS CURVAS FID/ $^1\text{H-RMN}$

Para estudiar la gran cantidad de información contenida en las curvas FID/ $^1\text{H-RMN}$ se aplicó el análisis de componentes principales (PCA). Las curvas FID/ $^1\text{H-RMN}$ corresponden a 150 muestras de PP en forma de pellets y los 200 muestras de CEP-LER en forma de polvo. Para el análisis de componentes principales fueron estudiadas todas las curvas FID/ $^1\text{H-RMN}$ adquiridas, incluyendo el conjunto completo de muestras en forma de pellets y polvo. El primer grupo analizado correspondió a las curvas FID/ $^1\text{H-RMN}$ de PP-h, CEP-LER, CEP-r y de CEP-i en forma de pellets. Los resultados obtenidos del PCA son mostrados en el gráfico de los loadings y los scores generados por los componentes principales (PC) de las curvas FID/ $^1\text{H-RMN}$. El PCA presenta; que el primer componente principal (PC1) explica un 96.3% de la variación en las curvas FID/ $^1\text{H-RMN}$. El segundo

(PC2) explica un 3.1% de la variación en las curvas FID/¹H-RMN y el tercero (PC3) explica tan solo un 0.5% de la variación de las curvas FID/¹H-RMN. Debido a que PC3 explica un bajo porcentaje de la variación, menor al 1%, la variabilidad del PC3 es considerada no significativa. Una representación gráfica de los scores para las diferentes curvas FID/¹H-RMN se presentan en la Figura 1, para PC1 Vs PC2 y PC1 Vs PC3.



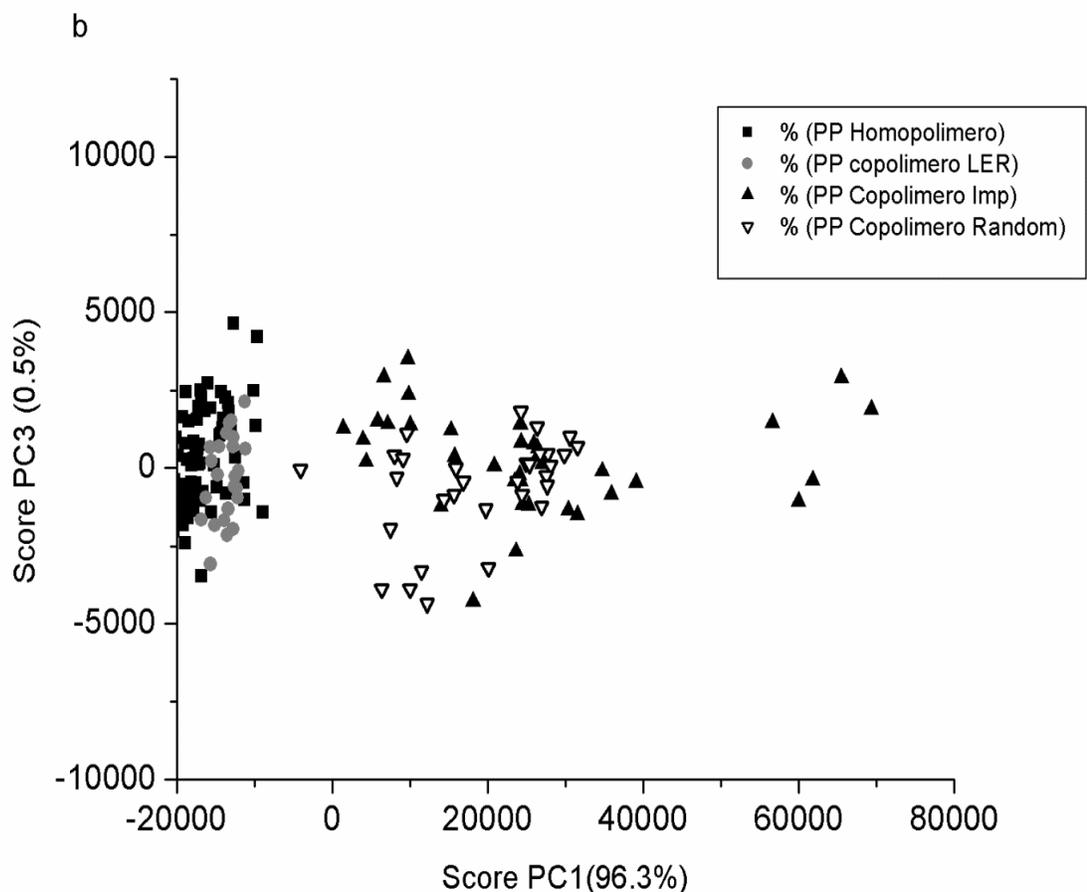


Figura 19. Representación gráfica de los scores obtenidos en el PCA para curvas FID de muestras de PP a) scores PC1 vs Score PC2 b) score PC1 vs score PC3.

4.3.1 INTERPRETACIÓN DE LOS SCORES DE PCA PARA LAS CURVAS FID/ ¹H-RMN

Los scores del PCA agrupan las muestras de acuerdo a su similitud formando diferentes clústeres o conglomerados. La Figura 19 a), presenta la distribución de las muestras de PP generada por el PC1 (96.3%) y el PC2 (3.1%). Los resultados muestran claramente 3 agrupamientos que diferencian las muestras de CEP-i, CEP-r y coloca en un solo grupo las

muestras de PP-h y CEP-LER produciendo un solapamiento explicado por la similitud estructural de estas dos familias de polipropileno. La distribución de las muestras obtenidas en el PC1 deja ver una posible relación inversa entre la cristalinidad de las muestras y la nueva variable obtenida, PC1 del análisis de PCA. Las muestras de PP-h y CEP-LER tienen alta cristalinidad por lo cual se agrupan de forma muy cercana, esto se debe a que el porcentaje de etileno de los polímeros LER es bastante pequeño (0.1 a 1.1%), mientras que los CEP-i (hasta 35% de etileno) y CEP-r (hasta 5.5% de etileno) forman grupos con tendencias cristalinas en direcciones opuestas, es decir mientras que los CEP-r tienden a organizar estructuras con grupos etilenos al azar, los CEP-i tienden a formar estructuras con cristales en bloque a partir de un núcleo de PP-h.

El componente principal PC2 describe un 3.1% de la variación en las curvas de FID/ ^1H -RMN en las muestras estudiadas. La distribución obtenida en el PC2 separa de forma clara las muestras de CEP-r y CEP-i orientando la tendencia de la cristalinidad de manera inversa. Los grupos PP-h, CEP-LER aparecen solapados en la distribución generada por el PC2.

El PC3 describe tan solo un 0.5% de la variación contenida en las curvas de FID/ ^1H -RMN. La representación gráfica de las muestras no presenta ningún clúster apreciable para las muestras de PP estudiadas. No es posible establecer a cual fuente de variación responde el PC3, por lo tanto se considera que no es significativo.

4.3.2 INTERPRETACIÓN DE LOS LOADINGS DEL PCA PARA LAS CURVAS

FID/ ^1H -RMN

Los loadings resultantes del PCA describen la variación espectral para las curvas FID de muestras de PP en pellets (Figura 20). El loading PC 1 (96.3% de la variación) tiene una forma similar a una Gaussiana de decaimiento rápido invertida (Figura 20(a)). La Gaussiana de decaimiento rápido es una curva asociada a la cristalinidad de la muestra, luego la relación entre cristalinidad de la muestra y el score PC1 es una relación inversa. El

resultado obtenido para el loading PC1 puede ser confirmado al comparar la distribución de las muestras en el PC score 1 y la cristalinidad asociada a cada muestra.

Teniendo en cuenta que para las muestras de PP estudiadas las condiciones físicas de cristalización (velocidad de enfriamiento) fueron constantes, es requerido buscar fuentes de variación para la cristalinidad obtenida en las muestras de PP estudiadas. El porcentaje de etileno es una de las variables de proceso utilizada para variar el grado de cristalinidad del PP. En este caso hay una variación general que coincide para cristalinidad y etileno. Pero factores adicionales como longitud de la cadena polimérica o aditivos usados en la preparación de las muestras de PP también podrían afectar la cristalinidad obtenida. El solapamiento de las muestras de PP-h y CEP-LER deja ver que el etileno no es la única variable responsable de las variaciones en cristalinidad que aparecen en el PC1.

El loading PC2 tiene una forma similar a la curva exponencial FID descrita en el análisis matemático de la curva (figura 20(b)). Teóricamente la curva exponencial FID está asociada a la estructura amorfa de las muestras estudiadas por $^1\text{H-RMN}$ (Ronald, Tache; Scott, Marino, Ajoy, Roy; Ronald, 1994). La disposición de las muestras en el score PC2 ubica en la parte superior del score a las muestras de CEP-r, en la parte intermedia las muestras de CEP-LER y PP-h, en la parte inferior del score están ubicadas las muestras de CEP-i. .

El loading PC3 tiene forma de gaussianan lenta. La variacion explicada por este PC es de 0.5% y no muestra una clara distribucion de las muestras en el caso de los scores. Es posible que en el caso de las muestras analizadas la variacion debida a las regiones de transicion cristalina/amorfa sea pobre para lograr clara difrencia entre las muestras.

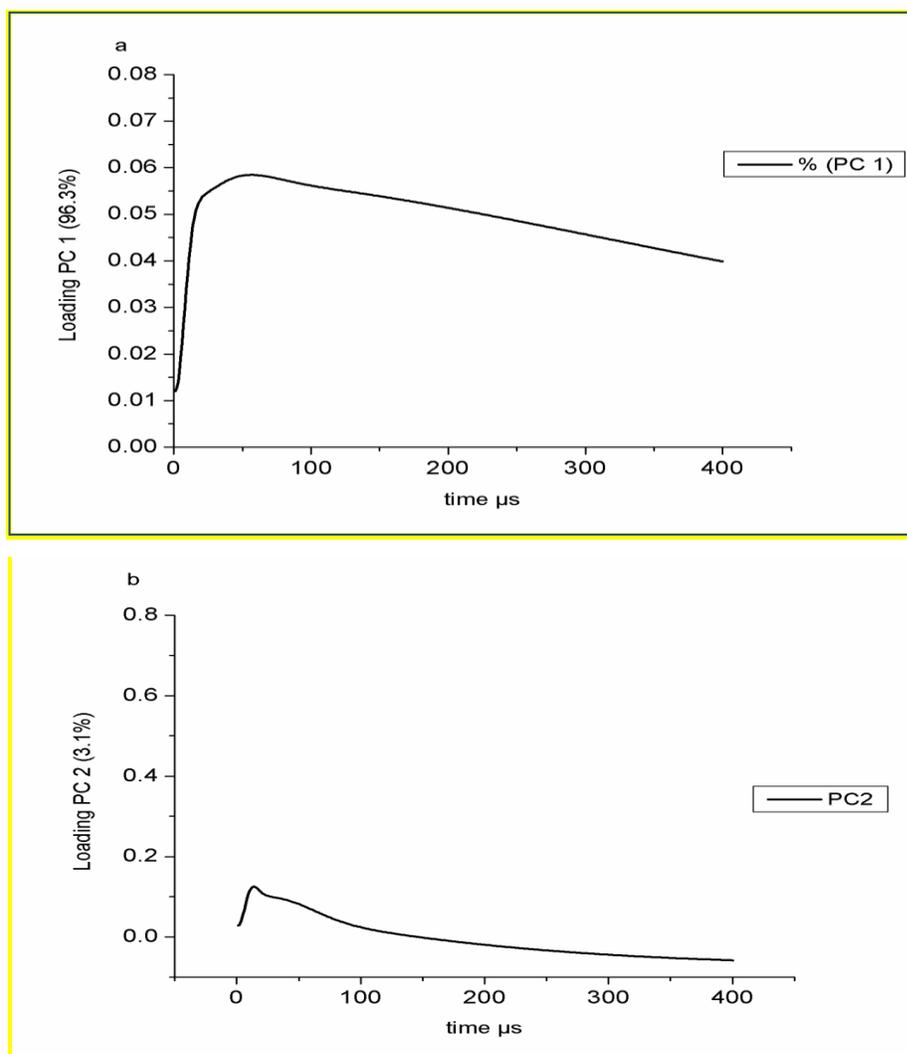


Figura 20. Representación gráfica de los loadings para el análisis PCA de curvas FID/¹H-RMN para muestras de PP en forma de pellets a) representación gráfica del loading PC1 b) representación gráfica del loading PC2.

Un segundo análisis de componentes principales fue aplicado para estudiar las diferencias entre las muestras CEP-LER en forma de pellets y polvo. Los resultados del análisis de componente principal muestran en el gráfico de score un PC1 con un 72.9% de la variación y un PC2 con el 24% de la variación. El score PC1 separa a las muestras en polvo y pellets (Figura 21). En el lado negativo del eje PC1 aparecen ubicadas las muestras en forma de polvo, en cambio en el lado positivo del eje PC1 aparecen las muestras en forma de pellets.

En el caso de los loading resultantes del PCA, la forma del loading PC1 es igual a la curva de decaimiento rápido asociada con la variación en cristalinidad de las muestras. Considerando la forma del PC1 y la distribución de las muestras, podemos afirmar que las muestras en forma de polvo poseen una menor estructura cristalina a la que presentan las muestras en forma de pellets; en este caso como ambas son muestras de copolímero LER, la composición química no es la fuente de variación, esto conduce que sean las condiciones físicas de enfriamiento (el historial de enfriamiento) de las muestras la principal fuente de variación con un 72.9% de la variación obtenida.

El segundo componente principal PC2 con un 24.4% de la variación espectral, separa las muestras según su porcentaje de estructura amorfa. El loading PC2 es similar a la curva exponencial FID asociada con la composición amorfa de las muestras estudiadas

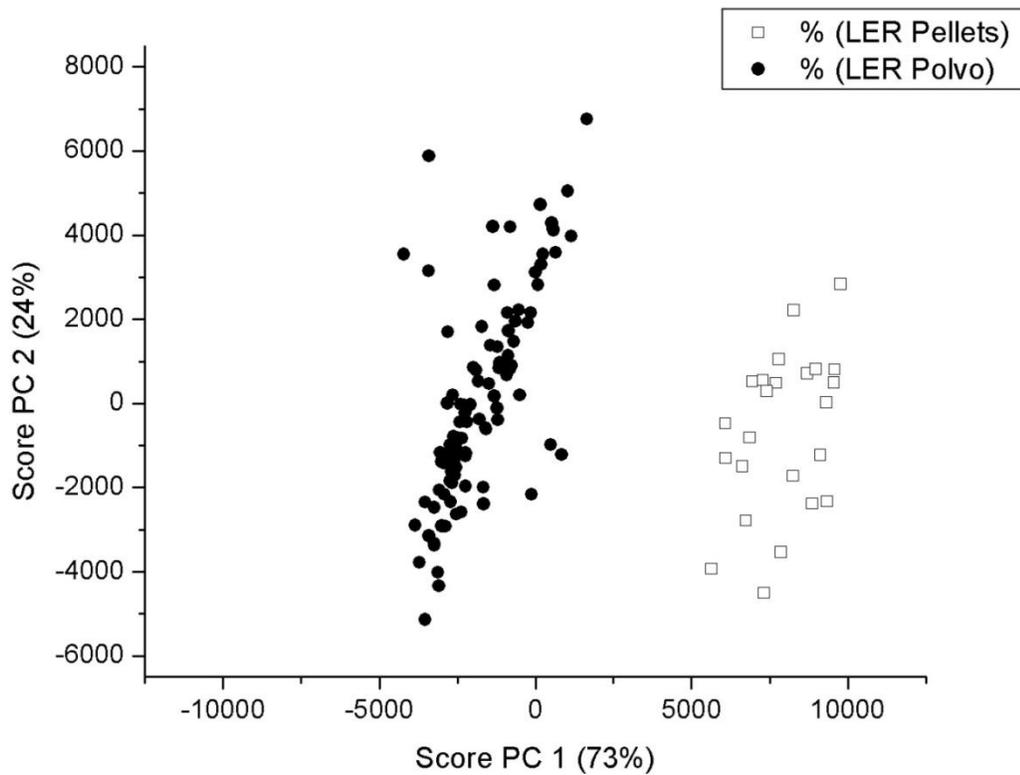


Figura 21. Representación gráfica de los escores obtenidos en el PCA para curvas FID de muestras de CPE-LER en polvo y pellets.

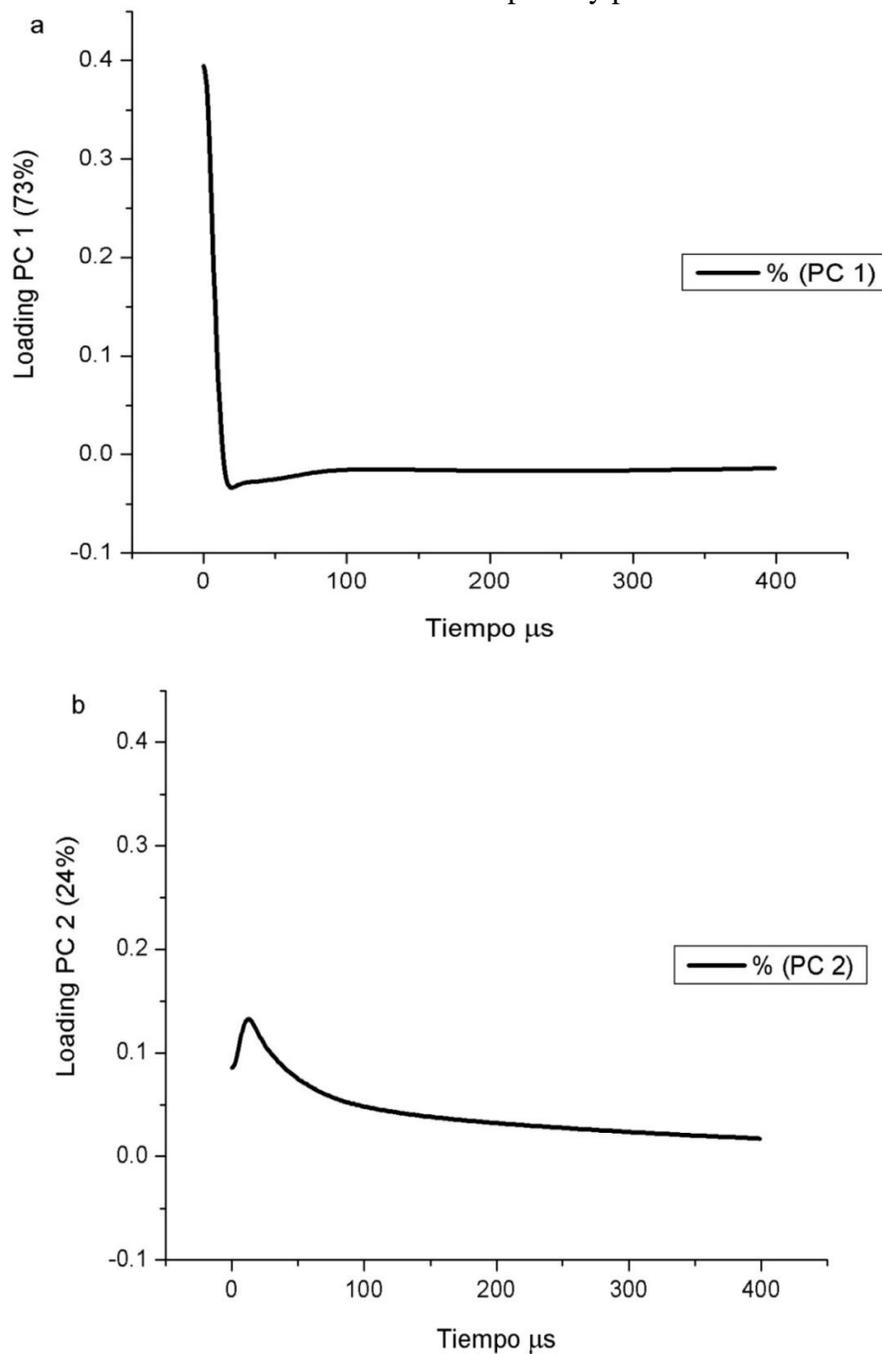


Figura 22. Representación gráfica de los loadings obtenidos en el PCA para curvas FID de muestras de CPE-LER en polvo y pellets a) Loading PC 1 b) loading PC 2.

4.4 PREDICCIÓN DE PROPIEDADES FISICOQUÍMICAS EN MUESTRAS DE POLIPROPILENO USANDO ESPECTROSCOPIA DE ¹H-RMN

4.4.1 Generación y selección de descriptores matemáticos para curvas FID ¹H-RMN

El software Scritp Generator 1.6 modeló las curvas de decaimiento FID/¹H-RMN para las 350 muestras de PP estudiadas. El software Scritp Generator 1.6 origina 6 índices matemáticos A , a , B , b , C y c . Las variables A , B , C son los valores de amplitud para la ecuación de FID que modelan el decaimiento ¹H-RMN (ecuación 135). Las variables a , b , y c son los valores de índice de decaimiento para cada una de las ecuaciones exponenciales que modelan las curvas FID ¹H-RMN de las muestra de PP estudiadas. Los 6 descriptores matemáticos de un grupo seleccionado de las muestras estudiadas aparecen registrados en la Tabla 7.

$$f(t) = Ae^{-(at^2)} + Be^{-(bt^2)} + Ce^{-(ct)} \quad (\text{EC. 134})$$

Tabla 7. Descriptores matemáticos primarios A , a , B , b , C y c para un grupo seleccionado de muestras estudiadas usando curvas FID generados por el software AzTec que controla y adquiere los datos del equipo de ¹H-RMN equipo.

Identificación de la Muestra	Familia	Presentación	A	a	B	b	C	C
PP-h MFR – 2	Homopolímero	Pellet	11142.37	9.57	2580.91	42.70	2090.65	230.06
PP-h MFR – 4	Homopolímero	Pellet	10291.49	9.74	2567.50	40.99	2122.85	172.51
PP-h MFR – 3	Homopolímero	Pellet	10879.49	9.63	2619.97	39.41	1893.98	156.25
PP-LER MFR – 3	LER	Pellet	9824.50	9.85	2578.89	42.32	2400.76	201.63
PP-LER MFR – 3	LER	Pellet	10310.60	9.87	2720.91	42.63	2584.92	203.75
PP-LER MFR – 3	LER	Pellet	10128.62	9.83	2642.50	42.32	2479.51	201.83
PP-i MFR – 60	Impacto	Pellet	10542.93	9.58	2461.16	42.25	2804.10	498.83
PP-i MFR – 60	Impacto	Pellet	11204.36	9.57	2565.80	42.58	3173.30	509.26
PP-i MFR – 5	Impacto	Pellet	8822.89	9.55	2223.05	41.70	4712.99	644.37
PP-r MFR – 7	Random	Pellet	8237.91	10.49	2663.51	50.01	5160.63	312.38
PP-r MFR – 7	Random	Pellet	9827.41	9.94	2664.14	44.37	2986.42	248.04
PP-r MFR – 1	Random	Pellet	8582.66	10.09	2620.45	47.78	3701.90	301.47
RXR 07a	LER	Polvo	6491.39	11.02	3234.21	43.22	2251.71	221.54
RXR 07b	LER	Polvo	6226.96	11.07	3158.96	44.11	2377.22	242.57
RXR 07c	LER	Polvo	6070.82	11.07	3086.06	44.30	2359.80	249.16

Luego del proceso de realización de los barridos RMN el software AzTec presenta la curva FID/¹H-RMN y genera los seis descriptores matemáticos primarios correspondientes a la curva obtenida.

Luego manualmente estos seis descriptores primarios fueron manualmente cargados a un software externo, llamado Script Generator para generar los 230 descriptores secundarios de cada muestra de PP. La Tabla 8 presenta un ejemplo de algunos de los descriptores complementarios empleados en el estudio de la información contenida en las curvas FID de las muestras de PP.

Tabla 8. Ejemplo de descriptores secundarios suplementarios generados por transformación así como combinaciones binarias, terciarias y cuaternarias de los descriptores primarios.

Combinaciones	Generación de Descriptores secundarios					
	1/a	1/c	1/b	1/A	1/C	1/B
	SQRT(A)	SQRT(C)	SQRT(B)	SQRT(a)	SQRT(c)	SQRT(b)
Primarias	1/SQRT(A)	1/SQRT(C)	1/SQRT(B)	1/SQRT(a)	1/SQRT(c)	1/SQRT(b)
	log(A)	log(C)	log(B)	log(a)	log(c)	log(b)
	c/a	b/a	b/c	a/B	c/B	b/B
	A*a	A*c	A*b	C*a	C*c	C*b
Secundarias	a/C	c/C	c/C	c/A	a/A	b/A
	B*a	B*c	B*b	a*a	c*c	b*b
	A/B*a	A/B*c	A/B*b	C/B*a	C/B*c	C/B*b
	C/A*a	C/A*c	C/A*b	B/A*a/c	B/A*a/b	B/A*c/b
Terciarias	B/C*a	B/C*c	B/C*b	A/C*b	A/B*C/B	A/B*a/c
	1/a*C/A	1/c*C/A	1/b*C/A	a/c*a	a/c*b	c/b*a
	A/C*A/B	A/C*a/c	A/C*a/b	A/C*c/b	A/C*a	A/C*c
Cuaternarias	C/B*a/c	C/B*a/b	C/B*c/b	a/c*a/b	A/C*A/C	C/B*C/B
	C/A*a/c	C/A*a/b	C/A*c/b	B/C*b/c	c/a*b/a	C/A*C/A
	B/C*a/c	B/C*a/b	B/C*c/b	C/B*c/a	C/B*b/a	C/B*b/c

El rango de variación de los descriptores suplementarios calculados varió en un rango muy amplio desde 0 hasta 7×10^6 . Al graficar los valores de los 230 descriptores generados fue posible observar la presencia de valores extremos que dificultan el análisis multivariado posterior (Figura 23(a)). Para eliminar una posible distorsión o enmascaramiento originada

por los valores extremos fue aplicada una transformación para escalar los datos (Figura 23(b)). Por consiguiente, los 230 descriptores secundarios fueron normalizados a 1, previo al análisis de componentes principales y la regresión por mínimos cuadrados parciales.

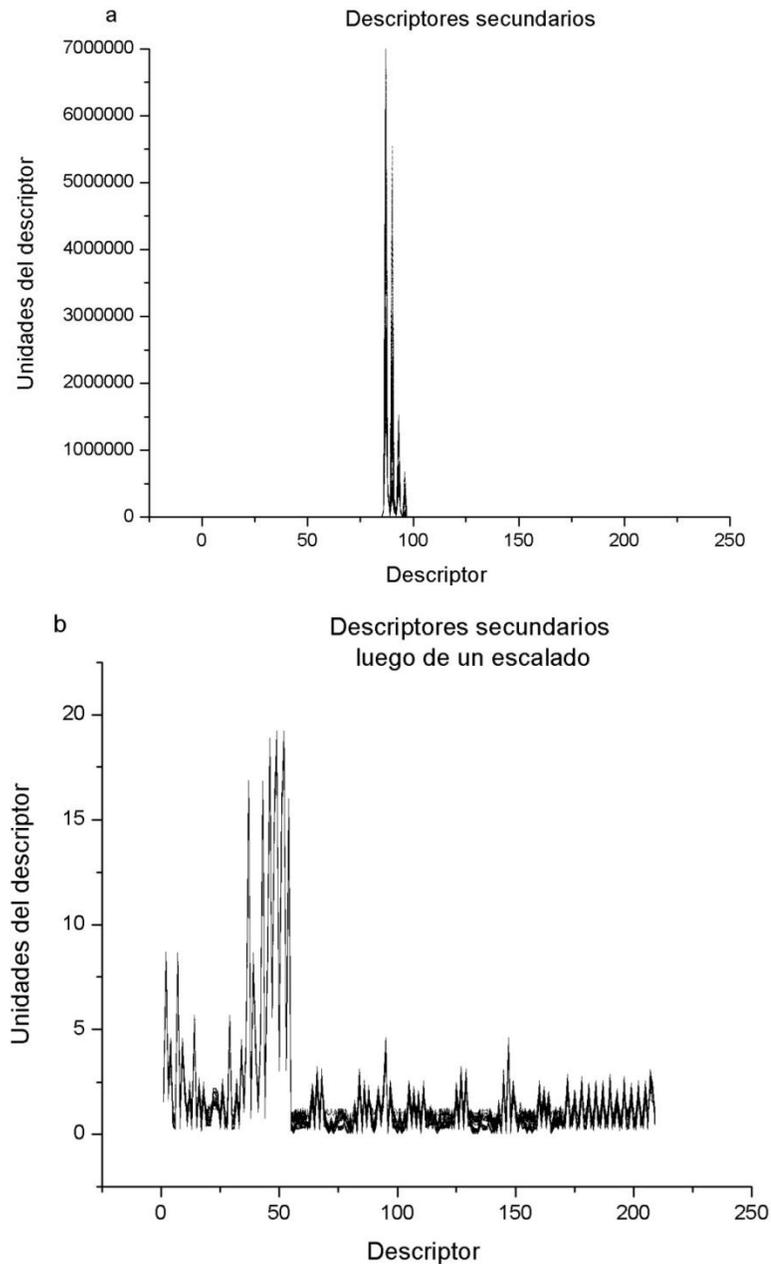


Figura 23. a) Representación gráfica de los descriptores secundarios obtenidos de la combinación de descriptores primarios b) Representación gráfica de los descriptores secundarios luego de un escalado.

4.4.2 ANÁLISIS DE COMPONENTE PRINCIPAL APLICADO A LOS DESCRIPTORES MATEMÁTICOS DE CURVAS FID/¹H-RMN

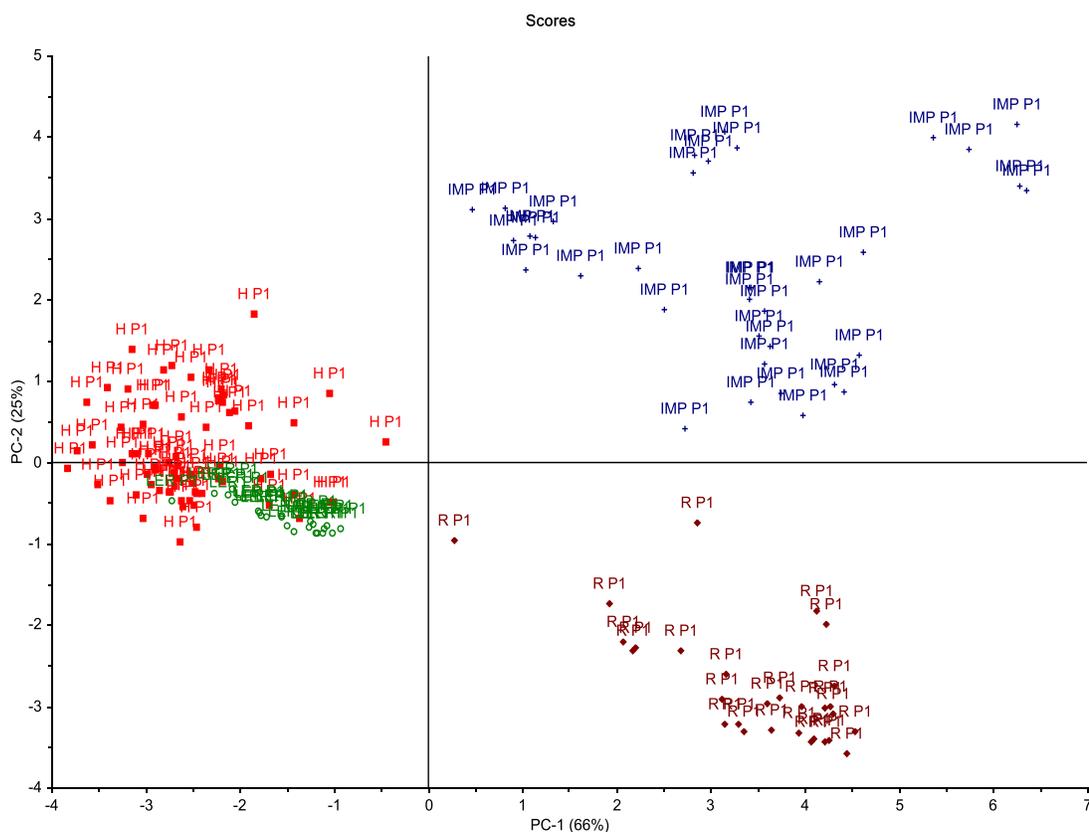


Figura 24. Representación gráfica del análisis de componente principal de los scores PC 1 vs PC 2

Para estudiar la información contenida en los descriptores matemáticos de las curvas FID/¹H-RMN, fue aplicado el análisis de componente principal (PCA). Los descriptores estudiados fueron los correspondientes a las 150 muestras de PP-h en forma de pellets y las 200 muestras de CEP-LER en forma de polvo.

El primer grupo analizado fue el de los descriptores matemáticos de las muestras pellets de PP-h, CEP-LER, CEP-r y CEP-i. El PCA de los scores presenta un primer componente

principal, PC1, que explica un 66%, el segundo PC2 explica un 25% y el tercer PC3 explica un 5% de la variación FID/¹H-RMN. Una representación gráfica de los scores es presentada en la Figura 24.

La Figura 24(a) presenta la distribución de muestras de PP generada por el PC1 (66%) y el PC2 (25%). Los resultados muestran la formación de clusters o agrupamientos de las muestras de PP-h y CEP-LER en la parte negativa del eje PC1. En la parte positiva del PC1 aparecen las muestras de CEP-r y CEP-i formando cada una un cluster.

4.4.3 INTERPRETACIÓN DE LOS PC LOADINGS PARA DESCRIPTORES MATEMÁTICOS DE CURVAS FID/¹H-RMN

Los resultados de los loadings en el PCA describen la variación de los descriptores matemáticos para las curvas FID de muestras de PP en forma de pellets (Figura 25). El loading PC 1 (66% de la variación) permita relacionar la variación obtenida en las muestras los descriptores. (Figura 25(a)).

Al comparar los resultados obtenidos usando PCA para las curvas FID/¹H-RMN y los resultados obtenidos para PCA usando descriptores matemáticos, es observado que los scores discriminan entre los cuatro diferentes tipos de PP que componen la muestra, este resultado es obtenido para ambos métodos de procesamiento de la señal. El análisis de los loading obtenidos del PCA son diferentes según el tipo de señal analizada. Los loadings derivados de las curvas FID/¹H-RMN están relacionados con la forma matemática de cada curva y la composición física de las muestras, (decaimiento rápido, decaimiento lento y exponencial). Los loadings obtenidos para el PCA realizado hecho a los descriptores matemáticos de las muestras, presentan colinearidad de la información.

En ambos casos, no existe una razón para seleccionar las curvas FID/¹H-RMN o los descriptores matemáticos para la generación de modelos de predicción PLS para las propiedades químicas y mecánicas de las muestras de polipropileno estudiadas.

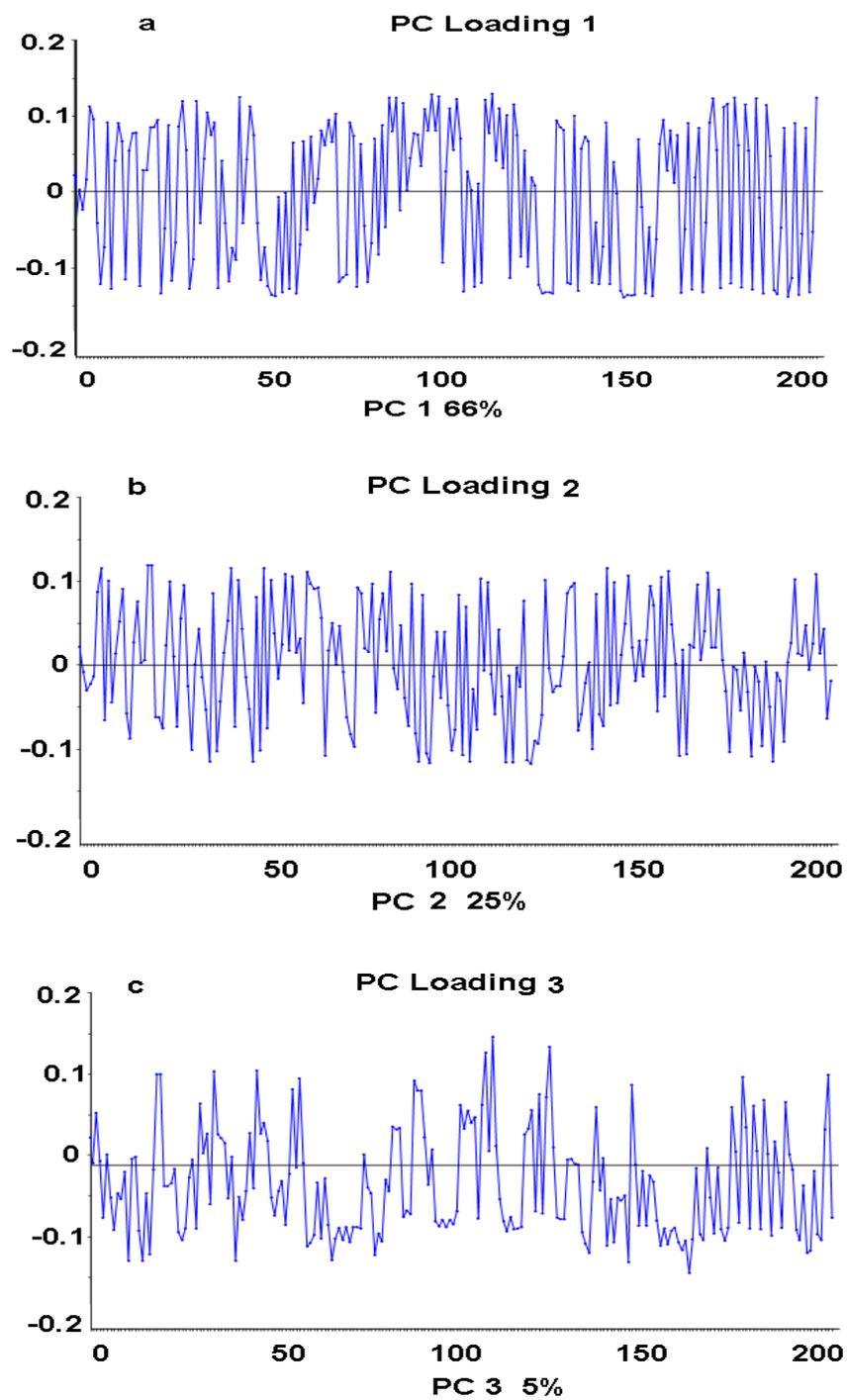


Figura 25. Representación gráfica de los loading generados por el PCA para descriptores matemáticos de curvas FID ^1H -RMN de muestras de PP. a) PC loading 1, b) PC loading 2, c) PC loading 3.

4.5 CALIBRACIÓN MULTIVARIADA CON REGRESIÓN POR MÍNIMOS CUADRADOS PARCIALES (PLS)

Uno de los objetivos de este trabajo es la cuantificación de propiedades fisicoquímicas y/o mecánicas de PP sólido usando espectroscopia $^1\text{H-RMN}$ y modelos de predicción multivariada PLS. La generación de los modelos de cuantificación para PP usando espectroscopia $^1\text{H-RMN}$, puede usar las curvas FID/ $^1\text{H-RMN}$ o los 230 descriptores matemáticos generados para cada curva. Estos fueron analizadas ambas opciones para determinar si existía diferencia significativa de los resultados obtenidos y a la vez determinar cuál tipo de modelo es recomendable aplicar $^1\text{H-RMN}$ para predecir propiedades fisicoquímicas/mecánicas del PP para el control de calidad de la producción industrial.

A continuación estudiaremos los resultados de la predicción de tres tipos de modelos PLS obtenidos para la predicción cuantitativa del contenido de etileno, resistencia a la tensión a la cedencia (Tensile strength at yield) y módulo de flexión 1%Secante, calculados para muestras sólidas de PP-h, CEP-LER y CEP-i.

4.6 PREDICCIÓN DE LA CONCENTRACIÓN DE ETILENO EN MUESTRAS DE CEP-LER USANDO ESPECTROSCOPIA $^1\text{H-RMN}$

Para predecir el contenido de etileno usando regresión por mínimos cuadrados parciales (PLS) y curvas FID/ $^1\text{H-RMN}$, fueron usados como datos de referencia los valores de etileno en muestras CEP-LER en polvo. Las muestras empleadas en el desarrollo de modelos PLS para la predicción del contenido de etileno fueron divididas aleatoriamente en dos grupos. Un grupo de calibración conformado por 160 muestras de CEP-LER y un grupo de validación de 70 muestras de CEP-LER en polvo. Las muestras en el grupo de calibración fueron las únicas incluidas en el desarrollo del modelo de regresión por PLS. Las muestras en el grupo de validación, fueron usadas para evaluar la efectividad y

robustez del modelo de PLS para predecir la concentración de etileno en el CEP-LER en polvo.

El contenido de etileno en los grupos de calibración y validación varían entre 0 y 1.1 % w/w. La concentración de etileno en las muestras de validación y calibración fueron obtenidas por espectroscopia de infrarrojo medio como método de referencia. Para el desarrollo de los modelos de predicción por PLS de contenido de etileno usando espectroscopia $^1\text{H-RMN}$ fueron usados por separado los datos de las curvas FID y los valores de los 230 descriptores matemáticos generados para cada muestra de CEP- LER.

El desarrollo de los modelos PLS de predicción de porcentaje de etileno se realizaron usando dos métodos: 1) El tiempo (μs) de las curvas FID/ $^1\text{H-RMN}$ experimentales obtenidas, se utilizó como variables de predicción y como variable de respuesta a las concentraciones de etileno medidas por Infrarrojo medio en el CEP-LER en polvo. 2) Los 230 descriptores matemáticos generados a partir de las combinaciones de los parámetros: A, a, B, b, C y c fueron usados como variables de predicción y de igual forma la variable de respuesta fueron las concentraciones de etileno medidas por Infrarrojo medio en el CEP-LER en polvo.

La Tabla 9 registra los resultados obtenidos para los diferentes modelos de predicción PLS de etileno usando los métodos usando las curvas FID/ $^1\text{H-RMN}$ y descriptores matemáticos.

Para la evaluación de los modelos creados, varios parámetros estadísticos como es el RMSEC (raíz del error cuadrático medio de calibración) el RMSEP (raíz del error cuadrático medio de predicción) bias (sesgo o parcialidad) y la correlación entre valor medido y valor predicho, fueron usados para medir el desempeño de los modelos PLS.

El resumen de la evaluación para un grupo de modelos PLS de predicción de etileno en CEP-LER aparece en la Tabla 9. Los modelos creados basados en curvas FID/ $^1\text{H-RMN}$

fueron escalados usando un pre-tratamiento denominado media centrada por columnas (substracción del promedio por columnas). Los modelos PLS basados en los descriptores matemáticos tienen como pre-tratamiento el auto-escalado de las variables. Adicional a los pre-tratamientos, cada modelo PLS creado tiene diferentes transformaciones quimiométricas. Las transformaciones empleadas, fueron SNV, primera derivada, segunda derivada, Log 10, GLS y no transformación. Otro importante parámetro para evaluar el desempeño de los modelos es el número de factores que requiere el modelo de calibración. Un número elevado de factores en el modelo puede causar modelos PLS sobrealimentados. Los modelos PLS sobrealimentados presentan errores altos en la predicción de muestras externas al modelo (muestras de validación).

Tabla 9. Evaluación de modelos PLS de predicción de contenido de etileno, basados en información de curvas FID/¹H-RMN y descriptores matemáticos de decaimiento magnético.

Variables de Predicción	Pre-tratamiento	Transformación	Número de Factores	R ² Calibración	Q ² Predicción	RMSEC	RMEV	%RSEV	Bias		
Curvas de FID/ ¹ H-RMN	Media centrada en columna	Ninguna	11	0.930919	0.940399	0.0642	0.0581	7.06	0.000		
		Log 10	11	0.914686	0.903039	0.0711	0.0739	8.97	0.000		
		SNV	11	0.865099	0.662704	0.0891	0.1768	21.4	0.027		
	Descriptores Matemáticos Pre-tratamiento.	Auto escalado en columna	Derivada 1 ⁰	11	0.930269	0.939936	0.0645	0.0583	7.08	0.010	
			Derivada 2 ⁰	11	0.933097	0.942269	0.0632	0.0576	6.99	0.011	
			GLS								
			Weighting (alpha 0.02)	2	0.931681	0.938786	0.0641	0.0598	7.26	0.014	
			Ninguna	9	0.949953	0.943181	0.0445	0.0505	6.13	-0.003	
			Log	9	0.478999	0.478999	0.1128	0.1315	15.97	-0.011	
			SNV	11	0.917486	0.908113	0.0573	0.0639	7.76	-0.012	
		Derivada 1 ⁰	5	0.947344	0.943928	0.0456	0.0499	6.06	-0.005		
		Derivada 2 ⁰	7	0.948778	0.944579	0.045	0.0501	6.09	-0.004		
		GLS									
		Weighting (alpha 0.02)	2	0.943032	0.940941	0.0473	0.0520	6.31	-0.005		

El número de factores del modelo seleccionados como validados dependen del tipo de transformación usada o pre-tratamiento previamente usado antes de desarrollar los modelos de PLS. Los pre-tratamientos usados en los modelos de predicción de la concentración de etileno fueron las transformaciones SNV, Log10, primera derivada, segunda derivada, y sin ninguna transformación. Para evaluar los modelos y determinar el número de óptimo de factores fue empleado el RMSEP (Error Cuadrático Medio de predicción) y el RESP (Error relativo de predicción). En el caso del modelo PLS-SNV basado en curvas FID/¹H-RMN, el RMSEP, RMSEC y RESP es 0.0891, 0.176 y 21.46%, respectivamente.

La selección del mejor modelo del modelo óptimo se limitada escoger el modelo más simple o con el menor número de factores usados para la predicción. En nuestro caso el mejor modelo obtenido fue utilizando el pre-tratamiento de GLS weighting con las variables de predicción con las curvas de FID/¹H-RMN. La transformación GLS Weighting, genera modelos con un menor requerimiento de factores (2 componentes para este caso). Los modelos PLS GLS

Weighting, sin embargo generan errores de predicción similares a los demás modelos creados pero se considera óptimo porque usa menos factores. Para el modelo PLS usando curvas de FID/¹H-RMN, el RMSEP fue de 0.06 unidades de concentración de etileno y un Bias de 0.014 unidades de concentración de etileno.

Por otro lado, la prueba F es usada para evaluar si los parámetros los parámetros de los modelos PLS-GLS FID/¹H-RMN y modelos PLS-GLS, no muestran diferencia significativa entre los resultados obtenidos mediante curvas FID/¹H-RMN y descriptores matemáticos, $F_{cal} = 1.312196 < F_{crit} = 1.936958$, $P = 0.23$. De este resultado es posible afirmar que los modelos de predicción de etileno usando curvas FID/¹H-RMN son equivalentes a los modelos PLS obtenidos usando descriptores matemáticos.

4.7 CALCULO DE LA EXACTITUD DEL MODELO PLS PARA PREDICCIÓN DE ETILENO

La exactitud del modelo de PLS con pre-tratamiento GLS Weighting (PLS-GLS/W) para las curvas de FID/¹H-RMN para predecir la concentración de etileno en CEP-LER en polvo es descrita mediante el RMSEP y por el Bias o sesgo se muestra en la Tabla 3. El Bias se toma como el promedio del error obtenido. En todos los modelos propuestos el Bias fue menor a 0.027 unidades de concentración de etileno (%w/w), en el modelo PLS GLS FID el bias fue de 0.014 unidades de concentración de etileno. La Figura 26 es una gráfica de control que muestra concordancia entre los resultados obtenidos, por el método de referencia IR medio y el modelo PLS GLS FID de predicción propuesto.

El modelo PLS-GLS/W aplicado a las curvas de FID/¹H-RMN permitió calcular el contenido de etileno en muestras de CEP-LER en polvo. Para establecer si las concentraciones de etileno en el CEP-LER en polvo basada en la prueba de espectroscopia de infrarrojo medio, considerada la como referencia, difiere o no con la nueva prueba propuesta utilizando las curvas de FID/¹H-RMN, se realizó una prueba t pareada, la cual mostro que no existe diferencia significativa entre los métodos citados ($t_{cal}=0.471 < t_{crit}=1.979$; $P=0.683$).

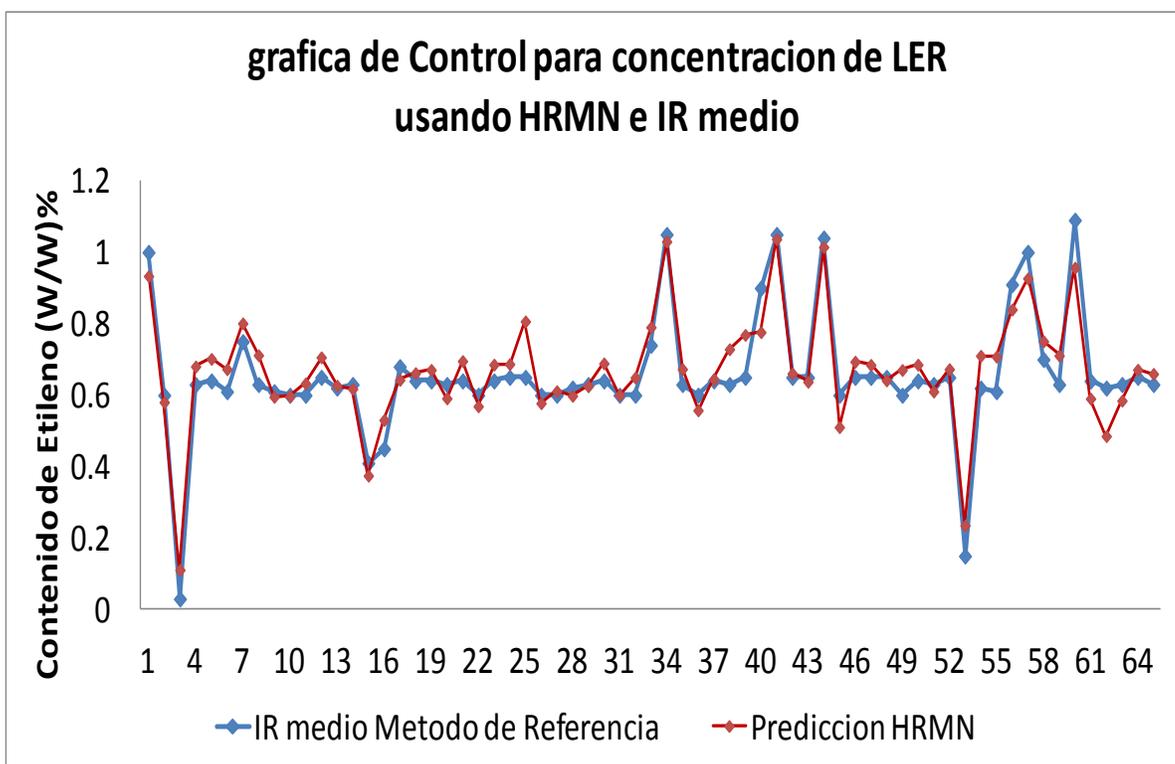


Figura 26. Predicción del modulo del contenido de etileno en muestras de polipropileno copolímero LER usando espectroscopia ^1H -RMN y comparando con lecturas por Infrarrojo medio de las mismas muestras diferentes a las usadas como calibración y validación.

La Figura 27 muestra la confrontación de los valores medidos por el método de referencia de IR medio y los obtenidos por el modelo PLS para todas las muestras. En una correlación entre los valores del método de referencia y el método propuesto, el coeficiente de correlación obtenido fue de 0.94, lo cual es una evidencia de una buena correlación entre los valores predichos por ^1H -RMN y los valores medidos de referencia IR medio.

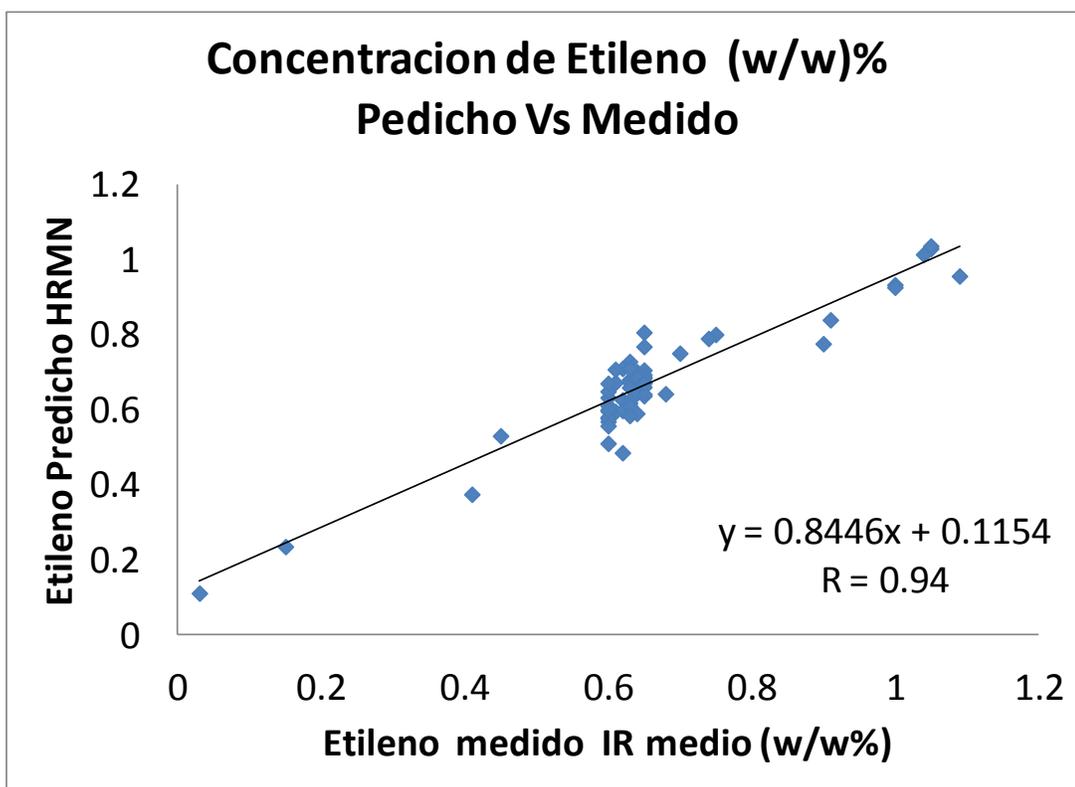


Figura 27. Correlación del % etileno en CEP-LER del método de referencia IR medio y del método PLS para FID/¹H-RMN

4.8 PREDICCIÓN DE LA RESISTENCIA A LA TENSIÓN PARA MUESTRAS DE PP-H, CEP-LER, CEP-R Y CEP-I USANDO CURVAS DE FID/¹H-RMN

En el desarrollo de modelos PLS para predecir la resistencia a la tensión, las muestras PP-h, CEP-LER, CEP-r y CEP-i en forma de pellets fueron clasificadas aleatoriamente en dos grupos: 1) Un grupo de calibración conformado por 120 muestras y 2) y el un grupo de validación de 30 muestras.

Para la predicción de la resistencia a la tensión usando regresión por mínimos cuadrados parciales (PLS) para las curvas FID/¹H-RMN, fueron usados como datos de referencia los valores de resistencia a la tensión en muestras PP-h, CEP-LER, CEP-r y CEP-i en forma

de pellets. Los valores de referencia de la resistencia a la tensión fueron medidos usando una máquina universal de ensayos (Instron, Norwood, MA). Las muestras en el grupo de calibración fueron las únicas incluidas en el desarrollo del modelo PLS-tensil. Las muestras en el grupo de validación, fueron usadas para evaluar la efectividad de los modelos PLS de predicción de la resistencia a la tensión de PP.

Tabla 8. Evaluación de modelos PLS-tensile de predicción de resistencia a la tensión, basados en información de curvas FID/¹H-RMN y descriptores matemáticos de decaimiento magnético.

Variables de Predicción	Pre-tratamiento	Transformación	Número de Factores	R ² Calibración	Q ² Validación	RMSE C	RMESV	RSEV %	Bias	
Tiempo (µS) de Curvas de FID/ ¹ H-RMN	Media centrada	None	3	0,97913	0,987876	187,067	135,388	1,53	9,52648	
		Log 10	3	0,98549	0,98635592	156,247	143,054	1,62	9,65062	
		SNV	6	0,9856	0,98842552	155,64	135,208	1,53	-5,99147	
		Derivada 1 ^o	5	0,98539	0,98870269	156,772	131,005	1,48	-8,00113	
		Derivada 2 ^o	10	0,98143	0,98239198	176,569	164,579	1,86	-11,2446	
		GLS								
		Weighting (alpha 0,02)	2	0,98857	0,98835267	138,74	136,314	1,54	31,1897	
		None	1	0,983442	0,97741546	165,2	183,434	2,07	-16,0433	
		Log 10	2	0,9861	0,98194094	151,48	165,073	1,87	-18,4833	
		SNV	3	0,60993	0,49057823	723,726	775,68	8,77	-180,642	
Descriptores Matemáticos Pre-tratamientos	Auto escalado	Derivada 1 ^o	-	-	-	-	-	-	-	
		Derivada 2 ^o	-	-	-	-	-	-	-	
		GLS								
		Weighting (alpha 0,02)	2	0,98533	0,97706909	155,595	184,656	2,09	-7,85125	

Para la resistencia a la tensión, los grupos de calibración y validación presentaron variación entre 2500-5500 psi. Para el desarrollo de los modelos de predicción PLS de la resistencia a

la tensión usando espectroscopia $^1\text{H-RMN}$ fueron usados por separado los datos de las curvas FID y los valores de los 230 descriptores matemáticos generados para cada muestra de PP estudiada. El resumen de la evaluación para los modelos PLS de predicción de resistencia a la tensión aparece en la Tabla 8. Los modelos PLS para la resistencia a la tensión (tensile) basados en curvas FID/ $^1\text{H-RMN}$ fueron escalados usando un pre-tratamiento dominado media centrada por columna (substracción del promedio por columnas). Los modelos PLS tensile basados en los descriptores matemáticos tienen como pre-tratamiento el auto-escalado de las variables. Adicional a los pre-tratamientos cada modelo PLS-tensile creado tiene diferentes transformaciones quimiométricas (SNV, primera derivada, segunda derivada, Log10, GLS y no transformación).

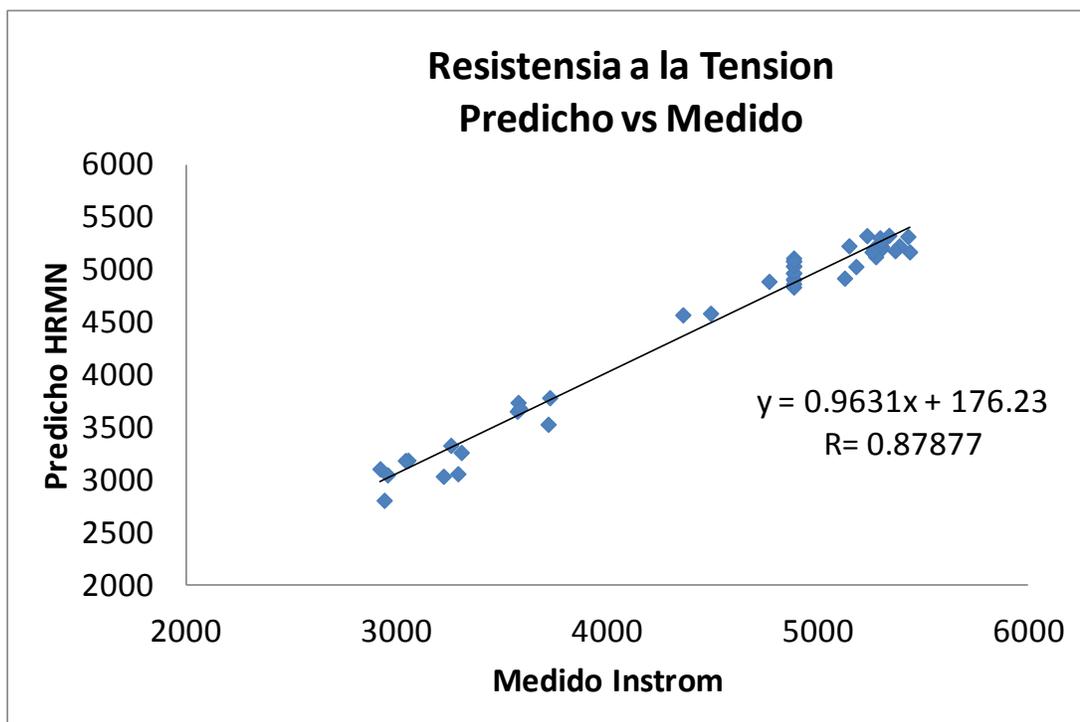


Figura 28. Valores de tensión medidos por el método de referencia y los valores predichos por el modelo PLS para las muestras de validación.

Las transformaciones quimiométricas usadas en los modelos PLS de predicción de resistencia a la tensión fueron: SNV, Log10, No transformación, primera derivada y segunda GLS. Para los modelos basados en los descriptores matemáticos las derivadas

carecen de interpretación y son descartados los modelos que las usan. GLS tiene valores de bias demasiado altos en ambos casos y se descartan los modelo PLS-GLS.

El modelo PLS Log10 para los descriptores tiene valor de Bias 10 veces mayor al promedio y por esto es descartado. El modelo PLS de segunda derivada es eliminado por requerir demasiados factores en comparación a los modelos restantes. El modelo seleccionado para la predicción de la resistencia a la tensión fue el PLS-Sin transformación FID/¹H-RMN. Este modelo no aplica transformación a los datos FID/H-RMN, usa 3 factores, tiene un error de predicción de 135 unidades de resistencia a la tensión y un error de predicción relativo de 1.53%. El modelo PLS-Sin transformación- DM tiene similares características, pero su error de predicción relativo es ligeramente mayor por ello se escoge el modelo PLS-Sin transformación FID/¹H-RMN.

Para este trabajo, la predicción de la resistencia a la tensión usando ¹H-RMN, fue realizada usando el modelo PLS-Sin transformación FID. Sin transformación, es equivalente a trabajar con los datos de la curva FID luego de mean center. El modelo PLS– Sin transformación FID/¹H-RMN, requiere 3 factores, con un RSEP de 1.5% y un bias de 9.5 psi. El modelo PLS– Sin transformación FID/¹H-RMN, generan errores de calibración de 187 psi errores de predicción de 135 psi.

4.9 EXACTITUD DE MODELOS DE PREDICCIÓN DE TENSILE PARA LAS DIFERENTES FAMILIAS DE PP

La exactitud del modelo PLS-tensile (Log10) FID/H-RMN para predecir la resistencia a la tensiones descrita mediante el RMSEP y por el bias tabla 8. En todos los modelos propuestos el bias fue menor a 180 psi para un rango de variación entre 2500-5500 psi, en el modelo seleccionado, PLS-tensile (Log10), el bias fue de 9.6 psi. La figura 29 muestra concordancia entre los resultados obtenidos, por el método de referencia Instron y el modelo PLS-tensile (Log10) de predicción propuesto.

Una prueba t pareada entre los resultados obtenidos por el método de referencia instron y las predicciones del modelo PLS-tensile (Log10) FID, no muestra diferencia significativa entre los métodos citados ($t_{\text{cal}}=0.471 < t_{\text{crit}}=1.979$; $P=0.683$).

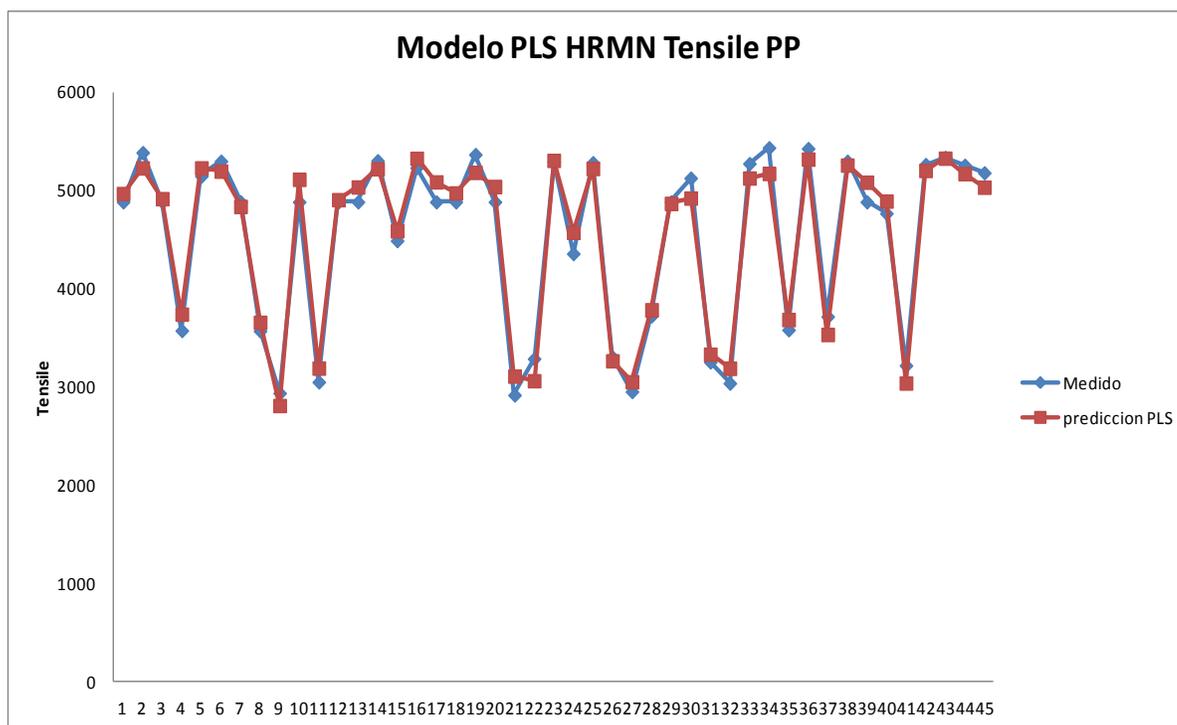


Figura 29. Comportamiento del valor medido y de la predicción de la resistencia a la tensión Secante para PP-h, copolímero LER, Random y de impacto usando espectroscopia $^1\text{H-RMN}$.

La Figura 29 muestra los valores medidos por el método de referencia Instron y los valores predichos por el modelo PLS para las muestras de validación. La correlación entre los valores de referencia y los predichos por el método PLS fue de 0.88 y la figura 26 son evidencia de una buena correlación entre los valores predichos por $^1\text{H-RMN}$ y los valores medidos de referencia para la resistencia a la tensión.

4.10 MODELO DE PREDICCIÓN DEL MÓDULO DE FLEXIÓN 1% SECANTE EN MUESTRAS DE POLIPROPILENO USANDO ESPECTROSCOPIA ¹H-RMN

En el desarrollo de modelos PLS para predecir el módulo de flexión, fueron usadas 160 muestras, clasificadas aleatoriamente en dos grupos. Un grupo de calibración conformado por 120 muestras de PP-h, CEP-LER, CEP-r y CEP-i en forma de pellets. Y un grupo de validación de 40 muestras de PP-h, CEP-LER, CEP-r y CEP-i en forma de pellets.

Para la predicción del módulo de flexión usando regresión por mínimos cuadrados parciales PLS y espectroscopia ¹H-RMN, fueron usados como datos de referencia los valores de módulo de flexión medidos con la máquina universal de ensayos Instron sobre probetas fabricadas por moldeo por inyección de las muestras de PP-h, CEP-LER, CEP-r y CEP-i. El modelo PLS fue obtenido con las muestras del grupo de calibración. Las muestras en el grupo de validación, fueron usadas para evaluar la efectividad de los modelos PLS de predicción del módulo de flexión de muestras de PP.

Los grupos de calibración y validación tienen un rango de variación entre 50000-270000 psi. Para el desarrollo de los modelos de predicción PLS de la resistencia del módulo de flexión 1% secante usando espectroscopia ¹H-RMN fueron creados datos de las curvas FID y valores de los 230 descriptores matemáticos generados para cada muestra de PP estudiada. Los datos de las curvas FID y de los descriptores matemáticos, fueron empleados para construir modelos de predicción PLS basados en curvas FID/¹H-RMN o modelos PLS basados en descriptores matemáticos.

El resumen de la evaluación para los modelos PLS de predicción para el módulo de flexión 1% secante aparece en la tabla 10. Los modelos PLS-Flexión basados en curvas FID/¹H-RMN fueron escalados usando un pre-tratamiento dominado media centrada por columnas. Los modelos PLS-Flexión basados en los descriptores matemáticos tienen como pre-tratamiento el auto-escalado de las variables. Adicional a los pre-tratamientos cada modelo PLS-flexión creados tienen diferentes transformaciones quimiométricas (SNV, primera derivada, segunda derivada, Log10, GLS y no transformación).

Las transformaciones quimiométricas usadas en los modelos PLS-Flexión de predicción del módulo de flexión fueron: SNV, Log10, No transformación, primera derivada, segunda derivada y GLS. Para los modelos basados en los descriptores matemáticos las derivadas carecen de interpretación y son descartados los modelos que las usan.

Los resultados muestran el modelo PLS-GLS DM como el mejor modelo para realizar la predicción del módulo de flexión 1% Secante. El modelo PLS-GLS DM tiene los valores de Bias más bajos (12.08). El modelo PLS-GLS DM requiere 2 factores para realizar la predicción con un error de predicción relativo de 5.33%. El otro modelo posible es el PLS-GLS FID/¹H-RMN con un error de predicción de 4.66%, pero con un requerimiento de 3 factores y un Bias de 982 psi. Los otros modelos de predicción PLS tienen valores de % RSEV y de Bias mayores al modelo PLS-GLS.

Para este trabajo, la predicción del módulo de flexión 1% Secante usando ¹H-RMN fue realizada aplicando el modelo PLS-GLS DM. El modelo PLS-GLS DM, requiere 2 factores, con un RSEP de 5.33% y un Bias de 12.08 psi. El modelo PLS-GLS DM, generan errores de calibración de 11425 psi y errores de predicción de 11068 psi.

4.11 EXACTITUD DEL MODELO DE PREDICCIÓN DEL MODELO DE FLEXIÓN 1% SECANTE PARA MUESTRAS DE POLIPROPILENO

La exactitud del modelo PLS GLS DM obtenido es descrita mediante el RMSEV y por el Bias Tabla 10. El Bias es el promedio del error obtenido. La figura 31 presenta la grafica de control de proceso para la evaluación del modulo de flexión 1% secante. La figura 31 muestra concordancia entre los resultados obtenidos, por el método de referencia Instron y el modelo PLS GLS de predicción propuesto.

El modelo PLS GLS calcula el módulo de flexión 1% Secante en muestras de PP en pellets directamente sin ningún pre-tratamiento. Una prueba t pareada entre los resultados obtenidos por el método de referencia Inyección de especímenes y luego medición con la máquina universal de ensayos Instron y las predicciones del modelo PLS ¹H-RMN, no muestra diferencia significativa entre los métodos citados ($t_{cal}=0.471 < t_{crit}=1.979$; $P=0.683$).

Tabla 10. Evaluación de modelos PLS Predicción módulo de Flexión 1% Secante, basados en información de curvas FID/¹H-RMN y descriptores matemáticos de decaimiento magnético.

	Pretratamiento	Mean center	R cal	R val	RMSEC	RMESV	RSEV %	Bias
	Transformación	No Descriptores						
FID/ ¹ H- RMN	None	4	0.976	0.979	11784.800	11617.300	5.60	615.853
	Log 10	2	0.977	0.978	11577.700	11833.000	5.70	1034.540
	SNV	6	0.978	0.982	11223.900	10746.300	5.18	1587.930
	Derivada 1	4	0.977	0.980	11595.000	11480.900	5.53	1821.880
	Derivada 2	8	0.977	0.978	11529.000	11973.300	5.77	1747.980
	GLS Weighting (alpha 0.02)	3	0.982	0.986	10273.500	9663.520	4.66	982.352
	Pretratamiento	Autoscale						
	Transformación	No Descriptores						
Descriptores Matemáticos	None	1	0.972	0.975	13465.8	11250.6	5.42	1086.07
	Log10	1	0.973	0.976	13278.1	11123.8	5.36	1288.71
	SNV	2	0.964	0.958	15212.2	14683.9	7.08	1539.4
	GLS Weighting (alpha 0.02)	2	0.980	0.976	11425.5	11068.9	5.33	12.0834

La Figura 30 muestra los valores medidos por el método tradicional de referencia ASTM y los valores predichos por el modelo PLS Log10 (FID/¹H-RMN) para las muestras de validación.

En la regresión de los datos el coeficiente de correlación obtenido fue de 0.978 y la figura 31 son evidencia de una buena correlación entre los valores predichos por ¹H-RMN y los valores medidos por el método de referencia ASTM.

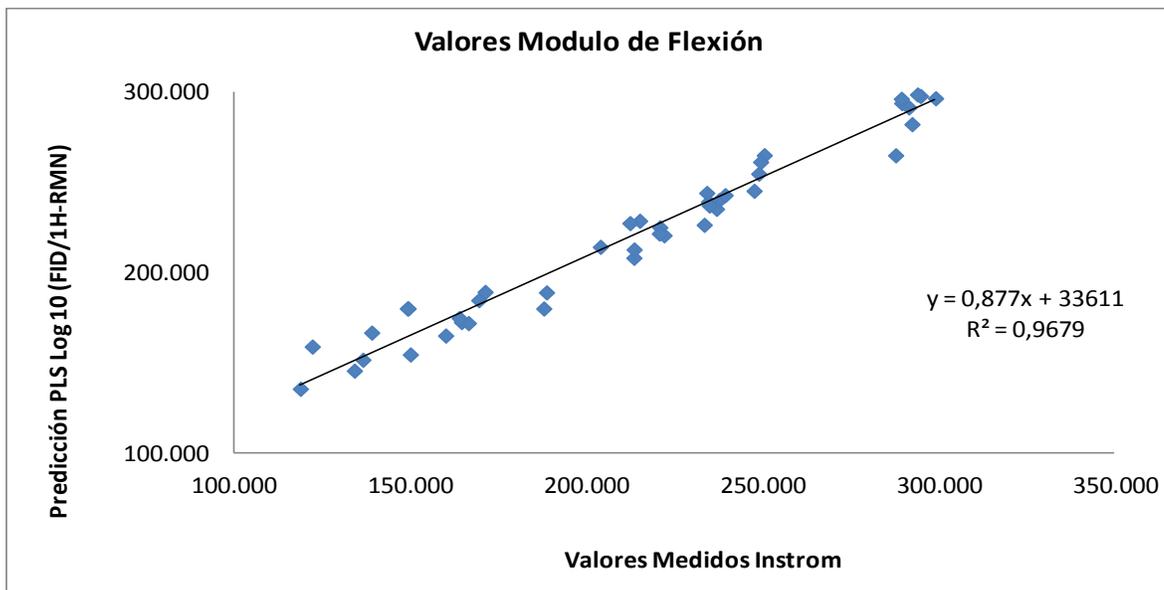


Figura 30. Modelo PLS log 10, empleando 2 factores en el modelo de predicción del módulo de flexión en pellets de PP.

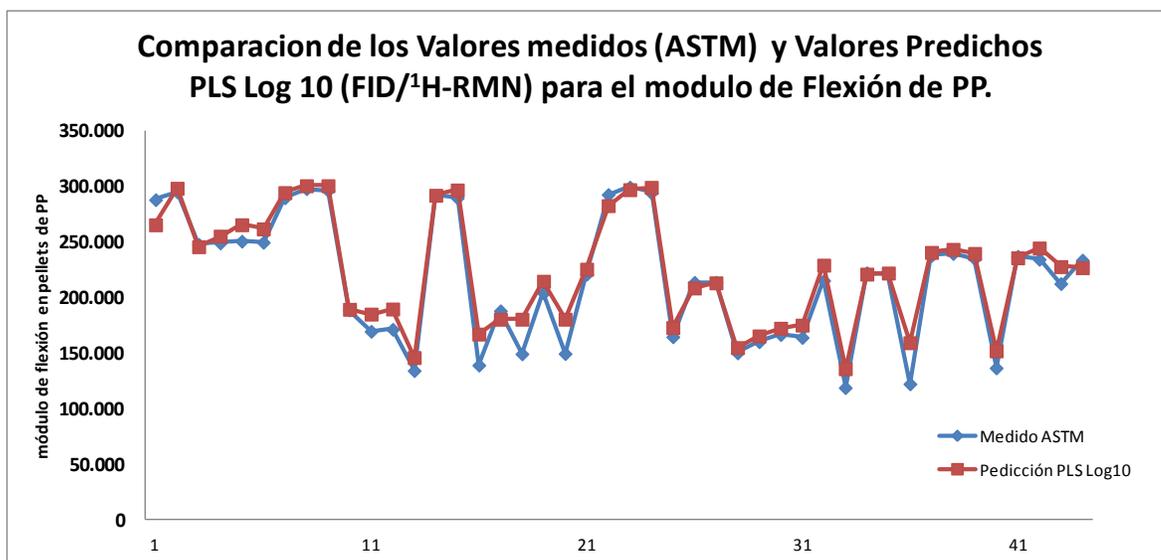


Figura 31. Comportamiento del valor medido y de la predicción del módulo de flexión 1% Secante para PP-h, copolímero LER, Random y de impacto usando espectroscopia ¹H-RMN.

5. CONCLUSIONES

Las curvas FID/¹H-RMN y los descriptores matemáticos para ¹H-RMN son adecuados para predecir el porcentaje de etileno en copolímeros LER, la resistencia a la tensión a la cedencia (Tensile Strength at Yield) y el módulo de flexión 1% secante en PP homopolímero, PP CPE-LER, PP copolímero random y PP copolímero de impacto. Teniendo en cuenta que las curvas FID/¹H-RMN y los descriptores son dos expresiones matemáticas del mismo análisis fisicoquímico, es posible afirmar que la información fisicoquímica de las curvas FID/¹H-RMN puede ser aprovechada usando directamente las curvas o, luego de la transformación de los descriptores matemáticos, sin que la decisión afecte de manera significativa los resultados obtenidos.

Las curvas FID pueden ser usadas para medir propiedades fisicoquímicas en diferentes partes del proceso de producción. La medición del contenido de etileno es realizada al material en forma de resina granular tomadas directamente del reactor de polimerización. Las mediciones de la resistencia a la tensión a la cedencia y del módulo de flexión 1% Secante son realizadas al producto terminado en forma de pellets.

Los modelos de predicción usando índices matemáticos o curvas FID/¹H-RMN son adecuados para extender su aplicación en la predicción del porcentaje de etileno, la resistencia a la tensión a la cedencia y el módulo de flexión 1% secante en el polipropileno usando equipos ¹H-RMN online, lo cual aumentaría la oportunidad de la información disponible para controlar el proceso de producción de polipropileno.

Es recomendable ampliar los rangos de variación de las propiedades fisicoquímicas estudiadas y aumentar el número de muestras incluidas en la calibración de los modelos PLS-FID/¹H-RMN.

6. BIBLIOGRAFÍA

- Arroyo, M., Lopez-Manchado, M. a., & Avalos, F. (1997). Crystallization kinetics of polypropylene: II. Effect of the addition of short glass fibres. *Polymer*, 38(22), 5587–5593.
- Bakeev, K. A. (2005). *Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*. (K. A. Bakeev, Ed.) (p. 451). Oxford: Blackwell Publishing Ltd.
- Bower, D. I. (2002). *An Introduction to Polymer Physics* (1st ed., p. 468). New York: Cambridge University Press.
- Bro, R. (2003). Multivariate calibration What is in chemometrics for the analytical chemist? *Analytica Chimica Acta*, 500(1-2), 185–194.
- Cabri, W., Ghetti, P., Pozzi, G., Alpegiani, M., Antibioticos, S. A., & Rivoltana, S. (2007). ReViews Polymorphisms and Patent , Market , and Legal Battles : Cefdinir Case Study Abstract : *Organic Process Research & Development*, 105459(September 1982).
- Candolfi, a, De Maesschalck, R., Massart, D. L., Hailey, P. a, & Harrington, a C. (1999). Identification of pharmaceutical excipients using NIR spectroscopy and SIMCA. *Journal of pharmaceutical and biomedical analysis*, 19(6), 923–35.
- Carraher, C. E. (2008). *Polymer Chemistry* (Seventh Ed., p. 721). Boca Raton, FL: Taylor & Francis Group, LLC.
- Choi, S.-G., & Kerr, W. L. (2003). ¹H NMR studies of molecular mobility in wheat starch. *Food Research International*, 36(4), 341–348.
- Derbyshire, W., Bosch, M. Van Den, Dusschoten, D. Van, Macnaughtan, W., Farhat, I. A., Hemminga, M. A., & Mitchell, J. R. (2004). Fitting of the beat pattern observed in NMR free-induction decay signals of concentrated carbohydrate – water solutions, 168, 278–283.
- Duer, M. J. (Ed.). (2002). *Solid-State NMR Spectroscopy Principles and Applications* (1 st., p. 563). Oxford: Blackwell Science Ltd.
- Edward P. Moore, J. (1996). *Polypropylene handbook*. (J. Edward P. Moore, Ed.) (p. 439).

- Fink, G., Steinmetz, B., Zechlin, J., Przybyla, C., & Tesche, B. (2000). Propene Polymerization with Silica-Supported Metallocene/MAO Catalysts. *Chemical reviews*, 100(4), 1377–90.
- Hansen, E. W., Kristiansen, P. E., & Pedersen, B. (1998). Crystallinity of Polyethylene Derived from Solid-State Proton NMR Free Induction Decay. *The Journal of Physical Chemistry B*, 102(28), 5444–5450.
- Harding, G. W., & Van Reenen, A. J. (2011). Polymerisation and structure–property relationships of Ziegler–Natta catalysed isotactic polypropylenes. *European Polymer Journal*, 47(1), 70–77.
- Karlstrom, H. (1995). Using multivariate methods on solid-state ^{13}C NMR data of complex materials. *Analytica Chimica Acta*, 315(1-2), 1–14.
- Keeler, J. (2002). *Understanding NMR Spectroscopy*. Wiley.
- Kristiansen, P. E., Hansen, E. W., & Pedersen, B. (2000). Phase distribution in polyethylene versus temperature probed by solid-state proton NMR free induction decay. *Polymer*, 41(1), 311–321.
- Litvinov, V. M. (2006). EPDM/PP thermoplastic vulcanizates as studied by proton NMR relaxation : Phase composition, molecular mobility, network structure in the rubbery phase, and network heterogeneity. *Macromolecules*, 39(25), 8727–8741.
- MACOMBER, R. S. (1998). *A Complete Introduction to Modern NMR Spectroscopy* (p. 400).
- McKenzie, J. S., Donarski, J. a., Wilson, J. C., & Charlton, A. J. (2011). Analysis of complex mixtures using high-resolution nuclear magnetic resonance spectroscopy and chemometrics. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 44.
- Mertens, B., Thompson, M., & Fearn, T. (1994). Principal component outlier detection and SIMCA: a synthesis. *The Analyst*, 119(12), 2777.
- Ning, N., Yin, Q., Luo, F., Zhang, Q., Du, R., & Fu, Q. (2007). Crystallization behavior and mechanical properties of polypropylene/halloysite composites. *Polymer*, 48(25), 7374–7384.
- Odian, G. (2004). *Principles of Polymerization*. *Journal of Controlled Release* (Vol. 22, p. 294). John Wiley & Sons, Inc.
- PACKER, V. J. M. and K. J. (1993). *NUCLEAR MAGNETIC RESONANCE IN SOLID POLYMERS* (1 st., p. 369). New York: CAMBRIDGE UNIVERSITY PRESS.

- Pandey, G. (2002). Nondestructive evaluation of physico-mechanical properties of polypropylene copolymer by infra-red spectroscopy. *European Polymer Journal*, 38(4), 745–750.
- Reich, G. (2005). Near-infrared spectroscopy and imaging: basic principles and pharmaceutical applications. *Advanced drug delivery reviews*, 57(8), 1109–43.
- Ronald, Tache; Scott, Marino, Ajoy, Roy; Ronald, D. (1994). US005302897A.
- Shlens, J. (2005). A Tutorial on Principal Component Analysis. *Measurement*.
- Tripathi, D. (2002). *Practical Guide to Polypropylene* (1 st., p. 103). Shawbury: Rapra Technology Limited.
- Vishu, S. (1984). *Handbook of Plastics Testing Technology*. New York: John Wiley & Son, Inc.
- Wise, B. (1996). The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*, 6(6), 329–348.
- Wold, S. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2), 109–130.
- Zhang, L., Hansen, E. W., Helland, I., Hinrichsen, E., Larsen, Å., & Roots, J. (2009). Crystallinity in Ethene–1-Hexene Copolymers Determined by ¹H and ¹³C NMR. A Comparative Study. *Macromolecules*, 42(14), 5189–5195.
- Zhang, P., Liu, X., & Li, Y. (2006). Influence of β -nucleating agent on the mechanics and crystallization characteristics of polypropylene. *Materials Science and Engineering: A*, 434(1-2), 310–313.