



COLEGIO DE POSTGRADUADOS

**INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN EN CIENCIAS
AGRÍCOLAS**

CAMPUS TABASCO

**PROGRAMA EN PRODUCCIÓN AGROALIMENTARIA EN EL
TRÓPICO**

**ESTUDIO DE LA DEGRADACIÓN DE HIDROCARBUROS EN
SUELOS MEDIANTE LA TÉCNICA DE INFRARROJO CERCANO**

VÍCTOR MANUEL VILLEGAS CORNELIO

T E S I S

**PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE**

MAESTRO EN CIENCIAS

H. CÁRDENAS, TABASCO


2010

La presente tesis titulada: **Estudio de la degradación de hidrocarburos en suelos mediante la técnica de infrarrojo cercano**, realizada por el alumno Víctor Manuel Villegas Cornelio, bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRO EN CIENCIAS
PRODUCCIÓN AGROALIMENTARIA EN EL TRÓPICO

CONSEJO PARTICULAR

CONSEJERO:


Dr. ARMANDO GUERRERO PEÑA

ASESOR:


Dr. JOEL ZAVALA CRUZ

ASESOR:


Dr. ALEJANDRO ALARCÓN

H. Cárdenas, Tabasco, 24 de noviembre de 2010

ESTUDIO DE LA DEGRADACIÓN DE HIDROCARBUROS EN SUELOS MEDIANTE LA TÉCNICA DE INFRARROJO CERCANO

Víctor Manuel Villegas Cornelio
Colegio de Postgraduados, 2010

Para el análisis de hidrocarburos en muestras de suelos contaminados existen diversos métodos, entre ellos, el más común es el Soxhlet, pero este implica etapas de pretratamiento, extracción, purificación-concentración y análisis. Sin embargo, estos procesos conllevan a la destrucción de las muestras, uso de reactivos y alto costo del análisis. No obstante, con el desarrollo de métodos instrumentales estas etapas se reducen, con la ventaja de disminuir al mínimo la manipulación de las muestras, el tiempo de análisis, el uso de reactivos químicos y las muestras no se destruyen en el análisis. La Espectroscopía de Infrarrojo Cercano (NIRS) como tecnología moderna proporciona datos sobre la naturaleza física y química de la muestra, con la generación de información contenida en las huellas espectrales y su posterior tratamiento por técnicas quimiométricas.

El objetivo del trabajo fue evaluar si NIRS genera información suficiente para diferenciar suelos contaminados con petróleo de aquellos que no lo están, así como para cuantificar los cambios en la degradación de petróleo crudo de una muestra contaminada. Para ello se utilizaron dos muestras de un suelo Gleysol. Una muestra fue contaminada con 50000 mg kg^{-1} de HTP y la segunda muestra no fue contaminada. Las huellas espectrales obtenidas con NIR en ambas muestras se trataron quimiométricamente, realizando primero el preprocesado mediante *Standard Normal Variate* (SNV) y una primera derivada, posteriormente se aplicaron técnicas para visualizar la distribución de los datos usando PCA. Seguido se generó el modelo de clasificación mediante LDA, el modelado de clases con SIMCA y la generación de un modelo de predicción por PLS. Los modelos se validaron interna y externamente con valores de poder de clasificación y predicción del 100%. Esto significa que los modelos generados son altamente efectivos para identificación de muestras contaminadas con petróleo crudo.

Para establecer modelos de estimación fue empleado el análisis por mínimos cuadrados parciales (PLS); los coeficientes de correlación múltiple (RSQ) y el error estándar de predicción corregido (SEP(C)) fueron: 0.98 y 13198.7 respectivamente para HTP. La capacidad de predicción del modelo (RPD) fue de 4.49. Por lo tanto, la tecnología NIR genera información suficiente para estimar la contaminación de los suelos con TPH, con resultados similares a los obtenidos con el método convencional de referencia (Soxhlet).

Palabras clave: Gleysol, contaminación, petróleo crudo, quimiometría, degradación de TPH, PCA, LDA, SIMCA, PLS.

STUDY OF THE DEGRADATION OF HYDROCARBONS IN SOIL BY NEAR INFRARED TECHNIQUE

Víctor Manuel Villegas Cornelio
Colegio de Postgraduados, 2010

For the analysis of hydrocarbons in contaminated soil samples there is a lot of methods, including the most common is the Soxhlet, which this involves stages of pretreatment, extraction, purification, concentration and analysis. However, the latest processes lead the sample destruction, utilization of special reagents, and increased cost of the analysis. Nevertheless, with the development of instrumental methods it is possible to reduce these stages, with the advantage of diminishing to a minimum sample handling, analysis time, and the use of chemical reagents, besides that samples are not destroyed throughout the analysis. Near Infrared Spectroscopy (NIRS) is a modern technology that provides data on the physical and chemical nature of the sample, with the generation of information contained in the spectral fingerprints and subsequent chemometric treatment.

The objective of this study was to evaluate whether NIRS generates enough information to distinguish oil-polluted soil from those that are not, and to quantify changes in the degradation of crude oil from a contaminated sample. Thus, two Gleysol soil samples were used, one sample was contaminated with 50000 mg kg⁻¹ of HTP and the second sample was kept uncontaminated. The NIR spectral fingerprints obtained in both soil samples were treated chemometrics, by first pre-processing using Standard Normal Variate (SNV) and first derivative, later techniques were applied to visualize the data distribution using PCA. Subsequently, we generated a classification model using PCA, modeling classes with SIMCA and generation a prediction model by PLS. The models were validated internally and externally with values of classification and prediction power of 100%. This means that the generated models are highly effective for the identification of samples contaminated with crude oil.

In order to establish estimation models an analysis by partial least squares (PLS) was

employed; the multiple correlation coefficients (RSQ) and the corrected prediction standard error (SEP(C)) were 0.98 and 13198.7, respectively, for HTP. The power of the predictive model (RPD) was 4.49. Therefore, NIR technology generates enough information to estimate the soil contamination with TPH, with similar results to those obtained with the conventional referential method (Soxhlet).

Key words: Gleysol, pollution, crude oil, chemometrics, degradation of TPH, PCA, LDA, SIMCA, PLS.

DEDICATORIA

A mi familia que me ha tenido gran paciencia y por su apoyo, gracias a ello pude lograr una meta más.

A mis hijos, la continuidad de mi ser. Siempre estarán en mi corazón.

A mi madre por todo su apoyo incondicional de siempre.

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por brindarme el apoyo económico para la realización de mis estudios de Maestría en el Colegio de Postgraduados.

Al Colegio de Postgraduados por darme la oportunidad de continuar con mi formación profesional en sus instalaciones y poner a mi disposición profesores e infraestructura.

A mi consejero Dr. Armando Guerrero Peña por su gran apoyo y paciencia en el desarrollo de la investigación, por los conocimientos transmitidos y por la preocupación que siempre mostró hacia mi persona.

Al Dr. Joel Zavala Cruz por los atinados consejos en la mejora de mi trabajo de investigación y por su disposición generosa cuando solicité su ayuda.

Al Dr. Alejandro Alarcón por su contribución al buen desarrollo y pertinentes comentarios hechos a este trabajo.

Al proyecto “Estudio Integral de la Cuenca Hidrológica del Río Tonalá y Lagunas del Carmen y Machona”, registrado en la Línea 1 Manejo Sustentable de los Recursos Naturales, del cual la presente tesis formó parte.

Al personal del Laboratorio de Análisis de Suelos, Plantas y Aguas (LASPA) por el apoyo en la realización de los análisis pertinentes de mi investigación.

A todos mis compañeros de generación, gracias por los buenos y difíciles momentos compartidos.

ÍNDICE

Índice de figuras	x
Índice de cuadros	ii
1. INTRODUCCIÓN GENERAL	1
1.1. Planteamiento del Problema	1
2. OBJETIVOS	3
2.1. Objetivo General	3
2.1.1. Objetivos específicos	3
3. HIPÓTESIS	4
3.1. Hipótesis General	4
3.1.1. Hipótesis específicas	4
4. REVISIÓN BIBLIOGRÁFICA	5
4.1. La Contaminación con Petróleo Crudo	5
4.2. Persistencia y Fraccionamiento de los Hidrocarburos del Petróleo en el Suelo	9
4.3. Normativa Ambiental	10
4.4. Composición Química y Propiedades Físicas del Petróleo Crudo	11
4.4.1. Hidrocarburos derivados del petróleo	12
4.4.1.1. Compuestos saturados e insaturados	13
4.4.1.2. Compuestos cíclicos	13
4.5. Técnicas para la Remediación del Suelo	14
4.6. Microorganismos Nativos en la Biodegradación	15
4.7. Técnicas para la Evaluación de Petróleo Crudo en Suelos	18
4.7.1. Métodos convencionales	18
4.7.2. Métodos modernos	19
4.7.2.1. Espectroscopia de infrarrojo cercano (NIRS)	20
4.8. Quimiometría	22
4.8.1. Reconocimiento de pautas no supervisadas	24
4.8.1.1. Análisis de agrupamientos	24
4.8.1.2. Análisis de componentes principales (PCA)	28
4.8.2. Reconocimiento de pautas supervisadas	31
4.8.2.1. Análisis discriminante lineal (LDA)	33
4.8.2.2. Modelo suave independiente de analogía de clases (SIMCA)	36
4.9. Modelos de Calibración	38
4.9.1. Calibración univariante	38
4.9.2. Regresión lineal múltiple	42
4.9.3. Regresión por mínimos cuadrados parciales (PLS)	43
4.9.3.1. PLS1	44
4.9.3.2. PLS2	45

4.9.3. Validación de modelos.....	46
4.10. Aplicación de los métodos quimiométricos	48
5. MATERIALES Y MÉTODOS	51
5.1. Etapa de Campo.....	51
5.1.1. Localización del área de estudio	51
5.1.2. Muestreo del suelo Gleysol.	52
5.1.3. Características del suelo Gleysol (GL)	54
5.2. Etapa de Laboratorio	54
5.2.1. Preparación de las muestras	54
5.2.2. Características físicas y químicas del suelo	55
5.2.2.1. Determinación de humedad.....	55
5.2.2.2. Determinación de pH	55
5.2.2.3. Determinación de materia orgánica	56
5.2.2.4. Determinación de textura	56
5.2.2.5. Determinación de conductividad eléctrica.....	56
5.2.2.6. Determinación de nitrógeno total	57
5.2.2.7. Determinación de capacidad de intercambio catiónico y bases intercambiables.....	57
5.3. Análisis de HTP Mediante el Método Soxhlet y la Tecnología NIR.....	58
5.3.1. Determinación de los TPH en muestras de suelo contaminado por el método convencional Soxhlet	58
5.3.2. Determinación de los TPH en muestras de suelo contaminado por espectroscopía de infrarrojo NIR	59
5.4. Análisis Quimiométrico de Resultados.....	60
5.5. Evaluación de la Población de Hongos Filamentosos Nativos	61
5.6. Características del Petróleo Crudo	61
5.7. Diseño Experimental.....	62
5.8. Establecimiento del Experimento.....	62
5.9. Análisis Estadístico de la Población de Hongos Filamentosos	63
6. RESULTADOS Y DISCUSIONES.....	64
6.1. Propiedades Físicas y Químicas del Suelo.....	64
6.1.1. Determinación de la humedad.....	64
6.1.2. Determinación del pH	66
6.1.3. Determinación de la materia orgánica	67
6.1.4. Determinación de la textura del suelo	69
6.1.5. Determinación de la conductividad eléctrica.....	70
6.1.6. Determinación del nitrógeno total	71
6.1.7. Determinación de la capacidad de intercambio catiónico y bases intercambiables.....	71
6.2. Estudio de la Población de Hongos Filamentosos	73
6.3. Análisis Químico de Hidrocarburos	75
6.3.1. Análisis mediante el método de referencia	75
6.4. Predicción de la Degradación del Petróleo Crudo en Suelo Gleysol	82
6.5. Interpretación de los Espectros NIR	84

6.5.1. Pretratamiento de los datos espectrales.....	90
6.5.1.1. Análisis de componentes principales (PCA)	91
6.6. Análisis Clasificadorio	96
6.6.1. Análisis discriminante lineal (LDA)	96
6.6.2. Modelado suave independiente de analogía clases (SIMCA)	99
6.6.3. Regresión por mínimos cuadrados parciales (PLS)	102
6.6.3.1. Ecuaciones de calibración	105
6.6.3.3.1. Validación de la ecuación de calibración	106
6.6.3.3.1.1. Validación interna	106
6.6.3.3.1.2. Validación externa	109
7. CONCLUSIONES	112
8. LITERATURA CITADA.....	114

ÍNDICE DE FIGURAS

Figura 1.	Localización de emergencias ambientales por derrame de petróleo en México	8
Figura 2.	Tipo de emergencia ambiental por derrame de petróleo en México	8
Figura 3.	Emergencias ambientales reportadas a la PROFEPA en el período 1993-2009	9
Figura 4.	Principales grupos de hidrocarburos que conforman el petróleo.....	13
Figura 5.	Técnicas fisicoquímicas y biológicas para la limpieza de suelos.....	15
Figura 6.	Distancia Euclídea y de Manhattan para las muestras k y l.	25
Figura 7.	Unión de un grupo con una muestra aislada de acuerdo con los métodos del centroide (a), de la unión simple (b) y de la unión completa (c).....	27
Figura 8.	Dendrograma que representa la cercanía entre muestras	28
Figura 9.	Análisis de componentes principales (PCA).....	29
Figura 10.	Representación de las puntuaciones de dos clases de objetos en los dos primeros componentes principales.....	31
Figura 11.	(a) Dos clases de objetos definidos por las variables x_1 y x_2 . (b) línea discriminante que separa las clases de objetos. (c) proyecciones de las muestras en la recta perpendicular a la línea discriminante y cálculo de centroides.....	34
Figura 12.	Distancia de Mahalanobis que determina la similitud entre dos variables multidimensionales.....	34
Figura 13.	Modelos de dos clases de objetos A y B en el espacio de sus componentes principales, regiones de confianza y predicción de dos muestras.....	37
Figura 14.	Diagrama de Coomans.....	38
Figura 15.	Matrices de PLS	45
Figura 16.	Ubicación Geográfica del sitio de muestreo del suelo Gleysol, en el Ejido Aquiles Serdán, La Venta, Tabasco	52
Figura 17.	Sitio de muestreo del suelo Gleysol (0-30 cm), Ejido Aquiles Serdán, La Venta, Tabasco.....	53
Figura 18.	Vegetación representativa del sitio de muestreo, Ejido Aquiles Serdán, La Venta, Tabasco.....	54
Figura 19.	Tren de extracción de HTP en equipo Soxhlet.....	58
Figura 20.	Instrumento NIR Versatile Agri Analyzer™ 6500.....	59
Figura 21.	Cuantificación de la población de hongos nativos en los tratamientos evaluados. Letras distintas en cada curva son estadísticamente diferentes (Tukey, $p=0.05$).....	74
Figura 22.	Concentración de petróleo crudo e hidrocarburos biogénicos por tiempo evaluado en suelo Gleysol.....	77
Figura 23.	Pérdida de HTP por tiempo en el tratamiento con suelo contaminado uno.....	82
Figura 24.	Línea de regresión y modelo de predicción para la degradación	

	de HTP y tiempo del bioensayos	84
Figura 25.	Espectros NIR característicos de las muestras de suelo no contaminado y contaminado con HTP, registrados al día uno de iniciado el experimento.....	86
Figura 26.	Interpretación de los espectros NIR de las muestras de suelo no contaminados y contaminados con HTP	87
Figura 27.	Espectros NIR de las muestras de suelo no contaminado y contaminado con HTP, a los 32 días de experimentación.....	89
Figura 28.	Proyección de las muestras de suelo en los dos primeros componentes principales.....	94
Figura 29.	Modelo de clasificación LDA, obtenido con los espectros NIRS de las muestras de suelo contaminado y no contaminado con petróleo crudo	98
Figura 30.	Diagrama de Coomans (SIMCA) obtenidos con los datos espectrales de NIR de suelo Gleysol no contaminado y contaminado con petróleo crudo	100
Figura 31.	Espectro NIR de las muestras de suelo no contaminado y contaminado con petróleo crudo del suelo Gleysol.....	104
Figura 32.	Modelo de calibración para HTP en suelo Gleysol no contaminado y contaminado con petróleo crudo	106
Figura 33.	Comparación de los valores de referencia con los valores predichos por el modelo PLS en NIR para la determinación de HTP en suelo no contaminado y contaminado con petróleo crudo	108
Figura 34.	Validación externa del modelo de calibración para HTP en suelo Gleysol no contaminado y contaminado con petróleo crudo	111

ÍNDICE DE CUADROS

Cuadro 1.	Número, localización y tipo de emergencia ambiental por derrame de petróleo en México	7
Cuadro 2.	Limites máximos permisibles para hidrocarburos específicos en suelos	11
Cuadro 3.	Propiedades del petróleo crudo “Olmeca”	62
Cuadro 4.	Descripción de tratamiento y unidades experimentales	62
Cuadro 5.	Propiedades físicas y químicas de la muestra de suelo Gleysol no contaminado	65
Cuadro 6.	Resultado de las determinaciones de Hidrocarburos Totales de Petróleo (HTP) por tiempo en el suelo Gleysol	81
Cuadro 7.	Degradación del petróleo crudo ligero tipo Olmeca en Gleysol con hongos nativos	82
Cuadro 8.	Poder discriminante de los Componentes Principales (CP) para las categorías suelo contaminado y suelo no contaminado	95
Cuadro 9.	Matrices LDA (de habilidad) de clasificación (a) y predicción (b) para las muestras de suelos contaminados y no contaminados con petróleo crudo	97
Cuadro 10.	Resultados SIMCA. Porcentaje de correcta clasificación (a) y predicción (b) para el suelo Gleysol contaminado y no contaminado	101
Cuadro 11.	Estimadores quimiométricos para la validación interna para HTP en suelo no contaminado y contaminado con petróleo crudo	108
Cuadro 12.	Estimadores quimiométricos para la validación externa para HTP en suelo no contaminado y contaminado con petróleo crudo	109
Cuadro 13.	Validación externa del modelo de calibración mediante la prueba de t de Student	111

1. INTRODUCCION GENERAL

1.1 Planteamiento del Problema

Los derrames de petróleo crudo causan grandes daños a la vida marina y terrestre, a la salud humana y a los recursos naturales. Por la tanto la caracterización adecuada de crudos derramados y la relación de éstos con las fuentes conocidas son extremadamente importantes para valorar los daños ambientales. Esta caracterización permitirá entender el destino y comportamiento así como la predicción del impacto potencial a largo plazo de los crudos derramados en el ambiente, seleccionar las respuestas adecuadas al derrame y tomar medidas efectivas de limpieza (Wang y Fingas, 2003).

Sin embargo, los procedimientos convencionales utilizados para resolver la mayoría de los problemas analíticos implican obtener información sobre los compuestos individuales de una mezcla determinada. El tratamiento de la muestra incluye, generalmente, etapas de separación y es la parte más costosa de todo el proceso. Además suele ser, también, la etapa más larga y una de las mayores fuentes de error. El procesado de los datos suele ser sencillo y la información obtenida se utiliza para detectar, identificar y cuantificar los analitos (Guerrero, 2006).

En algunas ocasiones no es necesario disponer de información individual sobre los compuestos de una muestra. El problema planteado puede resolverse satisfactoriamente con ciertas señales de perfil del conjunto de un grupo de compuestos cuando el problema no es qué contiene o no contiene la muestra sino determinar su identidad. La información contenida en los datos se extrae mediante la aplicación de técnicas quimiométricas. En este caso la manipulación de la muestra se reduce considerablemente, con lo que disminuyen los errores asociados a esta etapa y el tiempo total de análisis es muy corto al no ser necesarios procesos de tratamiento de la muestra. Ambas características ponen de manifiesto que esta

estrategia de procesado matemático de la información sobre el conjunto de compuestos sea muy atractiva.

El desarrollo de metodologías que permitan la reducción de la etapa de preparación de muestra es, por tanto, una de las tendencias actuales en química analítica. El objetivo es obtener la información necesaria para resolver el problema en el menor tiempo posible y reducir el coste del análisis.

En el presente trabajo se evaluó la capacidad de modelos generados con métodos modernos NIR y quimiometría para cuantificar la degradación de petróleo crudo de un suelo contaminado, comparado con el método convencional de extracción por Soxhlet.

2. OBJETIVOS

2.1. Objetivo General

Estudiar la degradación del petróleo crudo mediante espectrofotómetro de infrarrojo cercano y el método convencional Soxhlet.

2.1.1. Objetivos específicos

2.1.1.1. Generar reglas de clasificación para la identificación de muestras contaminadas y no contaminadas mediante espectrofotómetro de infrarrojo cercano.

2.1.1.2. Cuantificar los cambios en la concentración de petróleo crudo durante el proceso de degradación mediante espectrofotómetro de infrarrojo cercano.

3. HIPÓTESIS

3.1. Hipótesis General

La espectroscopía de infrarrojo cercano produce información similar al método convencional Soxhlet para estudiar la degradación de petróleo crudo en suelos.

3.1.1. Hipótesis específicas

3.1.1.1. El análisis por espectroscopía de infrarrojo cercano genera información suficiente y confiable para generar reglas de clasificación para la identificación de muestras contaminadas y no contaminadas.

3.1.1.2. El método por espectroscopía de infrarrojo cercano puede ser una alternativa metodológica a los métodos convencionales para cuantificar la concentración de petróleo crudo en suelos durante el proceso de degradación.

4. REVISIÓN BIBLIOGRÁFICA

4.1. La Contaminación con Petróleo Crudo

Las estadísticas mundiales del periodo 1978-1997, indican que ocurrieron en promedio, 289 derrames de petróleo crudo por año. De ellos, aproximadamente el 74 % de los accidentes corresponden a cantidades entre 40000-400000 L derramados, 22 % comprendido entre 400000 y 4000000 L, 3 % con 4000000-40000000 L y finalmente 1 % con más de 40000000 L derramados (Schmidt, 1999). En los últimos tiempos parece existir una tendencia a la disminución; los accidentes (derrames) disminuyeron un 228 % desde 1970 y un 154 % desde 1980 (Schmidt, 2001).

En México, la distribución de algunos de estos productos se realiza por una red de ductos (70%) que recorre gran parte del territorio nacional, conectando las diferentes estaciones y subestaciones de distribución y venta con las refinerías y plantas petroquímicas del país. Adicionalmente, se utiliza también la red de carreteras nacionales (30%) para el transporte de gasolina, diesel y algunos otros derivados del petróleo, lo cual implica un riesgo constante durante el traslado de estas sustancias (PROFEPA, 2010; Flores *et al.*, 2004) (Cuadro 1).

La contaminación de suelos con hidrocarburos es un problema que se ha vuelto muy común en nuestros días, debido principalmente al extensivo uso y consumo de combustibles derivados del petróleo a todo lo largo del país. En el estado de Tabasco, son frecuentes los derrames de petróleo en suelos. Estos derrames son originados por la rotura de oleoductos, fugas de hidrocarburos en presas de los pozos y descargas al medio ambiente en refinerías y otras instalaciones petroleras (Zavala *et al.*, 2007). Debido a que las acciones de restauración de suelos, por lo general no son inmediatas, los compuestos del petróleo pueden ser diseminados por el flujo de agua superficial en las llanuras de inundación, o infiltrarse al manto freático, generándose problemas de contaminación de suelo, agua, flora y fauna

(Palma-López *et al.*, 1999; Zavala *et al.*, 2007) e incluso la población que está en contacto con los sitios contaminados por hidrocarburos y sustancias químicas. De forma gráfica se presentan datos en porcentaje de localización y tipo de emergencia ambiental que se han presentado durante el periodo 1993-2009 en México (Figura 1 y 2).

Cuadro 1. Número, localización y tipo de emergencia ambiental por derrame de petróleo en México (PROFEPA, 2010).

Año	Número de eventos	Localización				Tipo							
		Terrestre		Marítima		Fuga o derrame		Explosión		Fuego		Otro	
		No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
1993	157	154	98.1	3	1.9	141	89.8	9	5.7	3	1.9	4	2.5
1994	416	389	93.5	27	6.5	359	86.3	21	5.0	28	6.7	8	1.9
1995	547	540	98.7	7	1.3	428	78.2	35	6.4	53	9.7	31	5.7
1996	587	578	98.5	9	1.5	460	78.4	34	5.8	70	11.9	23	3.9
1997	632	574	90.8	58	9.2	541	85.6	49	7.8	26	4.1	16	2.5
1998	566	507	89.6	59	10.4	502	88.7	30	5.3	29	5.1	5	0.9
1999	580	523	90.2	57	9.8	527	90.9	25	4.3	24	4.1	4	0.7
2000	596	552	92.6	44	7.4	529	88.8	26	4.4	35	5.9	6	1.0
2001	544	502	92.3	42	7.7	505	92.8	14	2.6	21	3.9	4	0.7
2002	473	438	92.6	35	7.4	426	90.1	16	3.4	28	5.9	3	0.6
2003	457	414	90.6	43	9.4	408	89.3	20	4.4	21	4.6	8	1.8
2004	503	390	77.5	113	22.5	474	94.2	10	2.0	19	3.8	0	0.0
2005	456	414	90.8	42	9.2	390	85.5	28	6.1	38	9.3	0	0.0
2006	362	349	96.4	13	3.6	302	83.4	31	8.6	29	8.0	0	0.0
2007	403	382	94.8	21	5.2	344	85.4	25	6.2	34	8.4	0	0.0
2008	349	343	98.3	6	1.7	302	86.5	16	4.6	30	8.6	1	0.3
2009	370	356	96.2	14	3.8	313	84.6	23	6.2	34	9.2	0	0.0
Total	7998	7405	92.6	593	7.4	6951	86.9	412	5.2	522	6.5	113	1.4

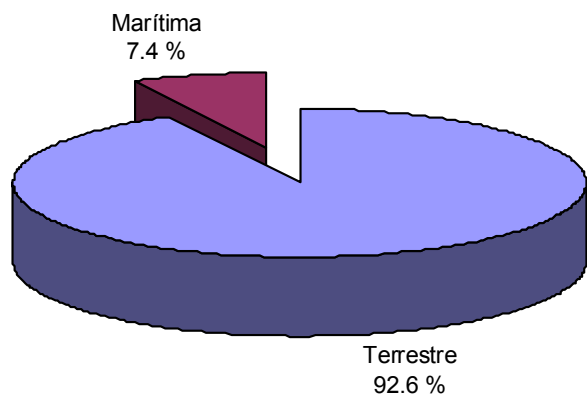


Figura 1. Localización de emergencias ambientales por derrame de petróleo en México (PROFEPA, 2010).

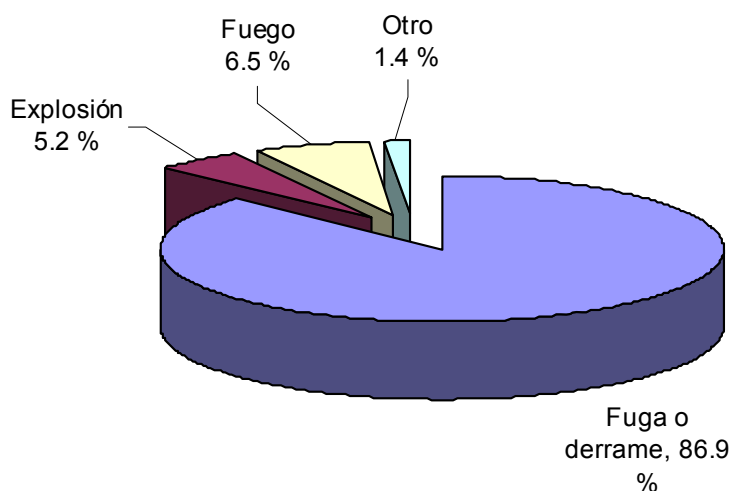


Figura 2. Tipo de emergencia ambiental por derrame de petróleo en México (PROFEPA, 2010).

El estado de Tabasco es el segundo productor de petróleo crudo en México con una producción de 452 859 barriles diarios (INEGI, 2002), sin embargo, tiene el segundo lugar a nivel nacional en emergencias ambientales reportadas, siendo de 1315 incidencias (PROFEPA, 2010) (Figura 3). De acuerdo con CIMADES (1999) se han registrado de 114-115 derrames de petróleo anuales, de los cuales 43 se presentan en el Activo Cinco Presidentes. En diagnósticos de suelos realizados en la región, fue posible cuantificar 100 g kg^{-1} en derrames recientes y 300 mg kg^{-1} en derrames antiguos (con más de un año), de hidrocarburos totales del petróleo, concentrados en una capa de aproximadamente 25 cm de espesor (Rivera-Cruz *et al.*, 2002). Por su

parte Zavala *et al.* (2005), cuantificaron para suelos Antrosoles, Acrisoles, Gleysoles e Histosoles, en el activo Cinco Presidentes 8569, 15324, 7782 y 15324 mg kg⁻¹ de HTP respectivamente, en los primeros 30 cm de profundidad del suelo, en derrames que van de 0.5 a 10 años de antigüedad.

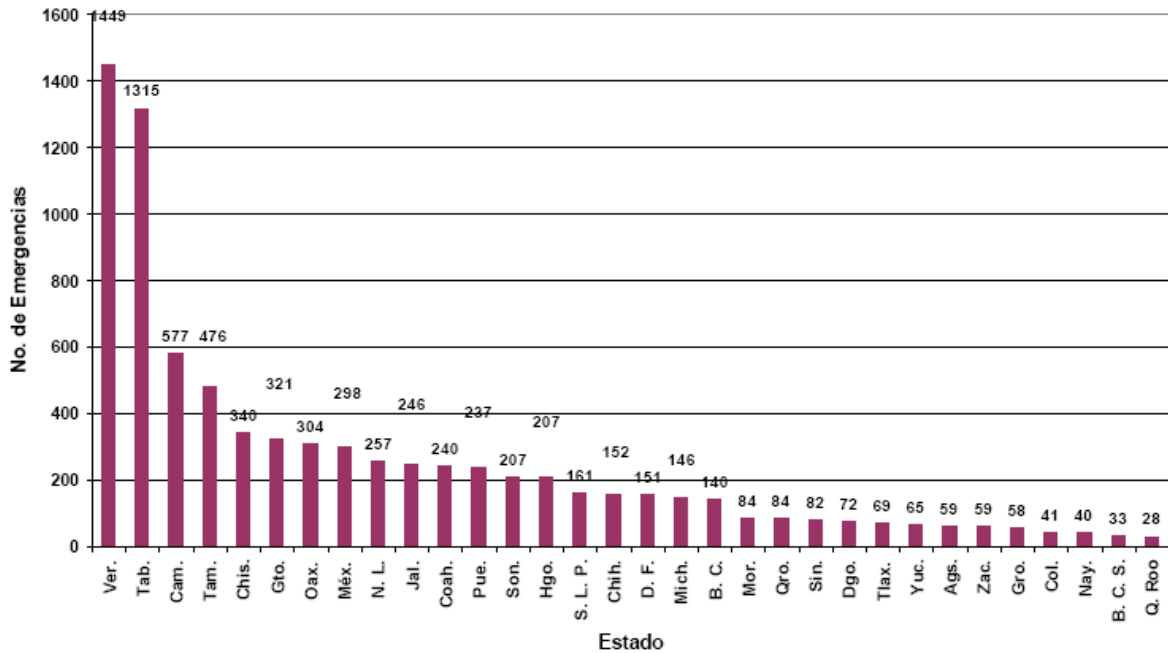


Figura 3. Emergencias ambientales reportadas a la PROFEPA en el período 1993-2009 (PROFEPA, 2010).

El origen de estos derrames se debe a la rotura de oleoductos, descontrol de pozos, descargas industriales, desbordamiento de presas de descargas de pozos e instalaciones de almacenamiento, separación y bombeo del petróleo (Botello, 1996; Adams *et al.*, 1999).

4.2. Persistencia y Fraccionamiento de los Hidrocarburos del Petróleo en el Suelo

Cuando el petróleo llega a la superficie del suelo por su alta viscosidad no penetra inicialmente hacia el subsuelo, se retiene en el horizonte superficial (PROFEPA, 2000). Una vez presente en el horizonte superficial, se separa en tres fracciones (Kesley *et al.*, 1997):

- a) Fracción volátil: Consiste de la fracción ligera que se fotooxida y volatiliza a través del espacio poroso del suelo y se transporta a la atmósfera, estos son los primeros cuatro n-alcanos (metano, etano, propano y butano).

- b) Fracción soluble: Esta fracción se difunde en la solución del suelo a través del flujo subterráneo de agua. Los hidrocarburos que entran en esta fase son de 5 a 17 carbonos, algunos ejemplos son: pentano, hexano, heptano, benceno, tolueno, etilbenceno, xileno y naftaleno.

- c) Fracción sólida: Es la fracción más estable del petróleo que queda adherida o adsorbida en la matriz del suelo; agrupa a los hidrocarburos con 18 o más átomos de carbono en su molécula. Incluye a los hidrocarburos aromáticos policíclicos (HAP) y a los compuestos polares.

4.3. Normativa Ambiental

La única norma en la que se especifican límites máximos permisibles para HAP individuales es la NOM-138-SEMARNAT-2003 (DOF, 2005) "que establece los límites máximos permisibles de hidrocarburos en suelos y especificaciones para su caracterización y restauración". Esta norma es específica para hidrocarburos del petróleo que provienen de derrames y/o fugas tanto recientes, como pasivos ambientales. En el cuadro 2 se indican los límites permisibles para los contaminantes altamente tóxicos.

Cuadro 2. Límites máximos permisibles para hidrocarburos específicos en suelos (DOF, 2005).

Hidrocarburos específicos	Uso de suelo predominante (mg kg ⁻¹ base seca)		
	Agrícola	Residencial	Industrial
Benceno	6	6	15
Tolueno	40	40	100
Etilbenceno	10	10	25
Xileno	40	40	100
Benzo(a)pireno	2	2	10
Dibenzo(a,b)antraceno	2	2	10
Benzo(a)antraceno	2	2	10
Benzo(b)fluoranteno	2	2	10
Benzo(k)fluoranteno	8	8	80
Indeno(1,2,3-cd)pireno	2	2	10

4.4. Composición Química y Propiedades Físicas del Petróleo Crudo

La industria mundial de hidrocarburos líquidos clasifica el petróleo de acuerdo a su densidad establecida como “Parámetro Internacional del Instituto Americano del Petróleo” (API) el cual diferencia los crudos por su calidad (IMP, 2008).

- Ligero. Tiene una densidad en el intervalo de 0.87-0.83 g/cm³ y en grados API en el intervalo de 31.1-39.
- Pesado. Tiene una densidad en el intervalo de 1.0-0.92 g/cm³ y en grados API en el intervalo de 10.0-22.3.
- Chapo. Tiene una densidad mayor a 1.0 g/cm³ y de 10.0 grados API.

La clasificación del petróleo se basa en sus propiedades y composición química (Dorn *et al.*, 1998):

- El petróleo ligero tiene altos contenidos de hidrocarburos saturados y contenidos bajos de compuestos polares.
- El petróleo pesado tiene menos hidrocarburos saturados y mayor concentración de compuestos polares, es por eso que el petróleo ligero es más tóxico que el pesado y se degrada más rápido.

- El petróleo, así como sus productos, son mezclas de hidrocarburos los cuales contienen principalmente hidrógeno y carbono en su estructura molecular.

El petróleo contiene gran variedad de hidrocarburos saturados, y sus productos tienen una amplia diversidad de usos como: gasolina, keroseno, solventes, petróleo crudo y parafinas, entre otros. Estos productos contienen mezclas de hidrocarburos que varían desde líquidos ligeros hasta sólidos. Los hidrocarburos de bajo peso molecular son gases, mientras que los de peso molecular más elevado, son líquidos o sólidos a temperatura ambiente y son insolubles en agua (Dorn *et al.*, 1998).

4.4.1. Hidrocarburos derivados del petróleo

Los hidrocarburos pueden ser saturados e insaturados, de cadena abierta (alifáticos) o cíclicos. En el grupo de los cíclicos, se encuentran los aromáticos, derivados del benceno que presentan resonancia (dobles enlaces no localizados). Los cíclicos no aromáticos presentan dobles o triples enlaces bien definidos y no tienen resonancia. Los enlaces covalentes entre los átomos de carbono de los hidrocarburos, pueden ser simples, dobles o triples (Figura 4), que tienen enlaces de carbono-carbono y carbono-hidrógeno, son polares e insolubles por lo que se disuelven bien en solventes no polares. Debido a su baja polaridad, los hidrocarburos que tienen de uno a cuatro carbonos por molécula, son generalmente gases a temperatura ambiente. Los hidrocarburos que tienen de 5 a 16 átomos, son líquidos a temperatura ambiente. Cuando los alcanos tienen 18 ó mas átomos de carbono, son parafinas sólidas a temperatura ambiente (Dorn *et al.*, 1998).

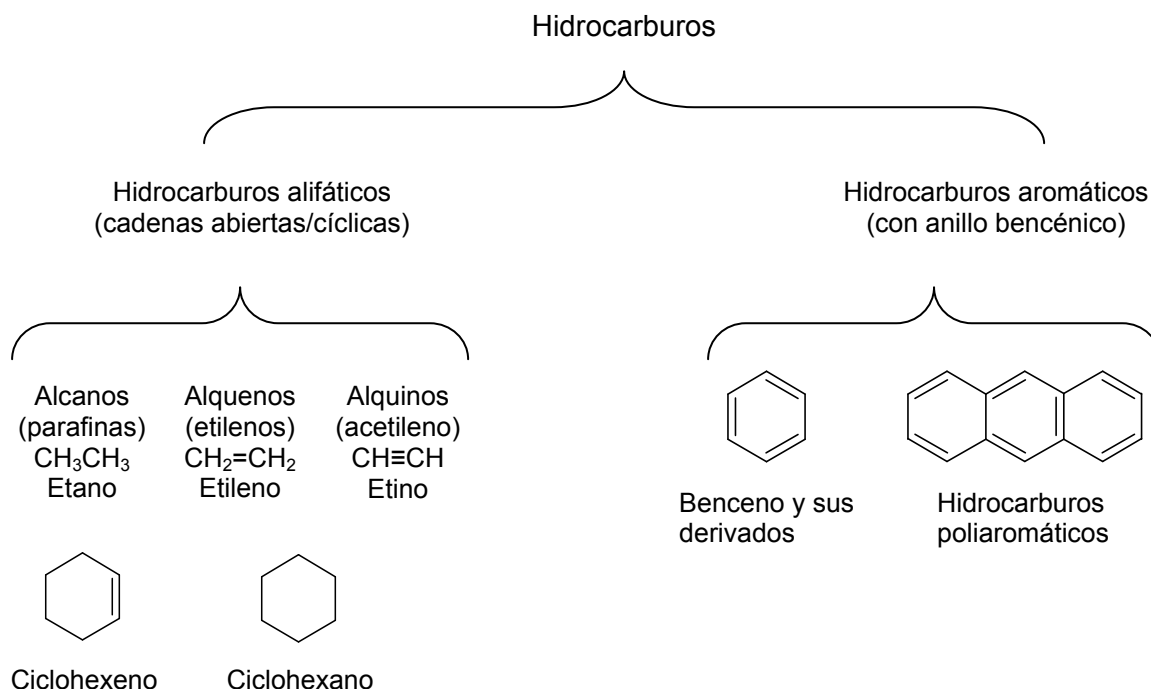


Figura 4. Principales grupos de hidrocarburos que conforman el petróleo (Dorn *et al.*, 1998).

4.4.1.1. Compuestos saturados e insaturados

Una forma sencilla de clasificar a los compuestos orgánicos se basa en la presencia o ausencia de sus enlaces. Si las moléculas de un compuesto incluyen enlaces sencillos se les denomina compuestos saturados o alcanos. Son hidrocarburos con moléculas de una sola ligadura (metano, etano, propano). En cambio los compuestos que contienen dobles o triples enlaces, se denominan compuestos insaturados, entre los que se encuentran a los alquenos y alquinos, los alquenos tienen doble ligadura y los alquinos tienen triple ligadura (Cuadrenys, 1981; Dorn *et al.*, 1998).

4.4.1.2 Compuestos cíclicos

Otro tipo de compuestos son los llamados hidrocarburos cíclicos que tienen algunos de sus átomos de carbono formando anillos, en el petróleo crudo hay una gran variedad de cicloalcanos o naftenos (Viñas, 2005). Estos compuestos tienen un arreglo de tres o

más átomos de carbono lo que forma un ciclo cerrado (ciclohexano, ciclohexeno, ciclohexino), a estos arreglos se les denomina compuestos cíclicos. Estos compuestos, pueden ser cicloalcanos, con enlaces simples, cicloalquenos con doble ligadura o cicloalquinos con triple ligadura. En el caso de los cicloalquinos, el anillo debe ser muy largo para tener un triple enlace, por lo que son raros de encontrar (Dorn *et al.*, 1998).

4.5 Técnicas para la Remediación del Suelo

La necesidad de remediar sitios dañados por la contaminación ha estimulado el desarrollo de nuevas tecnologías físicas, químicas y biológicas (Figura 5) que puedan ser capaces de reducir las concentraciones de hidrocarburos en los sitios impactados. Muchas de las tecnologías fisicoquímicas no destruyen totalmente el compuesto contaminante. Debido a esta limitante y al costo elevado de algunos procesos se están desarrollando y aplicando métodos innovadores de remediación (Fernández-Linares, 2000).

Las técnicas biológicas de limpieza hacen uso de sistemas biológicos, donde se aprovecha el potencial de los microorganismos para mineralizar o transformar contaminantes orgánicos en compuestos químicamente más sencillos, y con muy bajo impacto ambiental (Riser-Roberts, 1992). Las técnicas biológicas se han convertido en una alternativa prometedora para la recuperación de suelos contaminados con petróleo crudo (Atlas, 1981).

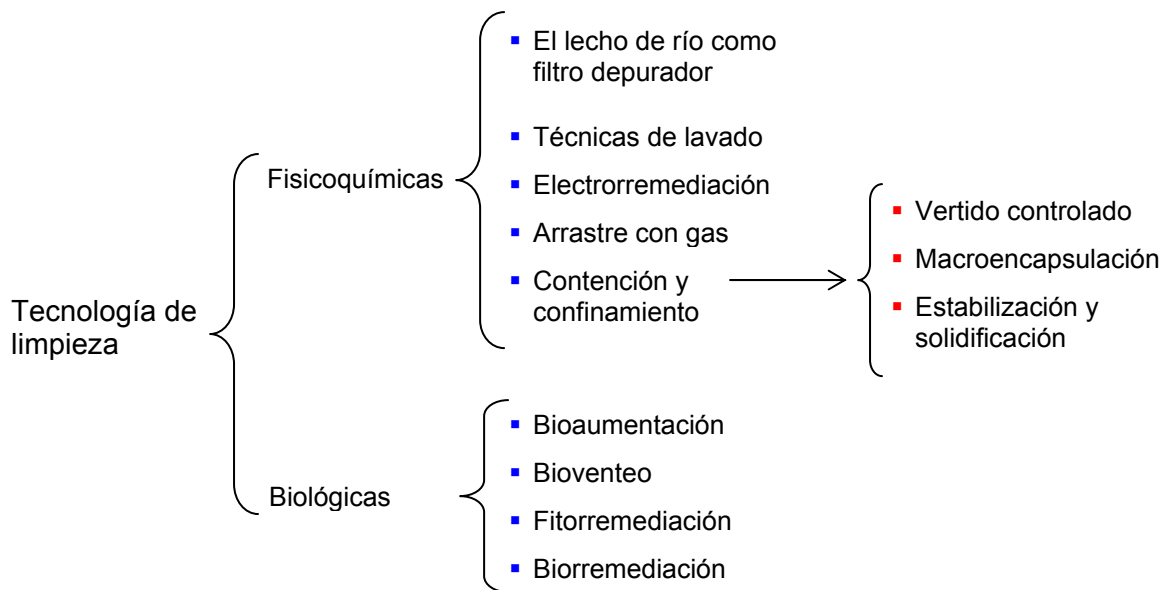


Figura 5. Técnicas fisicoquímicas y biológicas para la limpieza de suelos (Atlas, 1981).

4.6. Microorganismos Nativos en la Biodegradación

En la naturaleza los microorganismos son los principales agentes de la recirculación del carbono. Las comunidades naturales de microorganismos en estos hábitats tienen una versatilidad fisiológica asombrosa. Ellos son capaces de metabolizar y a menudo mineralizar una enorme cantidad de moléculas orgánicas. Probablemente todos los productos naturales, independientemente de su complejidad, es degradada por una u otra especie en algún entorno particular (Alexander, 1999).

En muchos ecosistemas ya hay una adecuada comunidad microbiana nativa hidrocarbonoclasta capaz de biodegradar una amplia variedad de hidrocarburos del petróleo, siempre que las condiciones ambientales sean favorables para la actividad metabólica degradante del petróleo (Atlas, 1977). Esta ha sido demostrada en el suelo, ambientes marinos y agua dulce. Algunos investigadores han sugerido que todos los suelos, excepto aquellos que son muy ácidos, contienen microorganismos capaces de degradar productos derivados del petróleo, en los cuales la inoculación microbiana no es necesario, en este caso es suministrar los nutrientes necesarios en el sitio contaminado (Riser-Roberts, 1992). La habilidad de utilizar los hidrocarburos esta

ampliamente distribuida entre poblaciones microbianas diversas (Atlas, 1981).

Muchas especies de bacterias, cianobacterias, hongos filamentosos y levaduras coexisten en los ecosistemas naturales y muchos actúan independientemente o en combinación para metabolizar hidrocarburos aromáticos (Cerniglia, 1984; Gibson, 1982). En general, el nivel de población que utilizan los hidrocarburos y su proporción dentro de la comunidad microbiana puede ser un índice sensible del ambiente a la exposición de hidrocarburos (Atlas, 1981). La degradación extensiva de contaminantes derivados del petróleo, por lo general, se realiza por mezcla de poblaciones microbianas, en lugar de una sola especie (Atlas, 1978).

Hay ventajas en el uso de los microorganismos nativos, en lugar de añadir microorganismos para degradar los desechos. A través de generaciones de evolución, las poblaciones naturales que se han desarrollado son ideales por su supervivencia y proliferación en el medio ambiente. Esto es particularmente cierto para sitios de desechos peligrosos no controlados donde los microorganismos han sido expuestos a los residuos por años o incluso décadas, para conseguir una ruta metabólica degradadora, donde se observa un enriquecimiento selectivo y cambios genéticos que causan una mayor proporción de microorganismos degradadores de hidrocarburos (Leahy y Colwell, 1990; Riser-Roberts, 1992; Ensley y Zylstra, 1997; Levin y Gealt, 1997).

Sin embargo, en muchas ocasiones, los microorganismos nativos requieren el aumento de nutrientes para lograr una degradación relativamente rápida y completa de los residuos peligrosos. Los organismos en condiciones naturales generalmente presentan carencias en fósforo, nitrógeno y azufre. La adición de estos nutrimentos estimula el crecimiento de la población natural para mejorar su metabolismo (Ensley y Zylstra, 1997). En el ambiente la microflora responde a la adición de petróleo o productos del petróleo. La corta persistencia de estos hidrocarburos indica que existen poblaciones vigorosas en el suelo. Sin embargo, la tasa de degradación es notablemente influenciada por la longitud de la cadena del hidrocarburo (Alexander, 1999). Este autor,

menciona también que aún cuando ciertas bacterias y hongos actúan sobre una amplia gama de compuestos orgánicos, no hay organismo conocido hasta la fecha capaz de destruir un gran porcentaje de la sustancia química.

De la variedad microbiana del suelo, las bacterias son las más empleadas en el proceso de biorremediación, aunque también se han empleado otros microorganismos como hongos, algas, cianobacterias y actinomicetos para la degradación de compuestos tóxicos en el suelo (Atlas y Bartha, 2002).

Una de las características sobresaliente de las bacterias hidrocarbonoclastas, es que poseen actividades peroxidasas y oxigenasas, que permiten la oxidación de algunas fracciones del petróleo (Heichen *et al.*, 2000). Esta oxidación cambia las propiedades de los compuestos haciéndolos susceptibles a ataques secundarios y facilitando su conversión a bióxido de carbono y agua (Van Hamme *et al.*, 2000; Torres, 2003).

En el caso de los hongos, se ha demostrado que son responsables del metabolismo de una gran variedad de compuestos como azúcares, proteínas e hidrocarburos biogénicos de origen vegetal o animal, así como sustancias procedentes del petróleo (Rivera, 2001). Los hongos oxidan a los hidrocarburos aromáticos mediante un sistema enzimático a base de citocromo P₄₅₀ monooxigenasa (Rosenberg y Ron, 1996). Boldu *et al.* (2002) estudiaron el papel del hongo *Cladophialophora sp.* sobre la degradación de benceno, tolueno, etilbenzeno y xileno, el cual degradó los compuestos alcalinizados (tolueno, etilbenzeno y xileno) pero no el benceno. El mecanismo de degradación es una combinación de asimilación y cometabolismo (García-Rivero y Peralta-Pérez, 2008). El tolueno y el etilbenzeno fueron usados como fuente de carbono y energía. En el proceso degradativo actúa la enzima monooxigenasa la cual se encargó de la degradación del tolueno, etilbenzeno y el xileno.

Para los hongos filamentosos se han propuesto modelos, para ubicar intracelularmente las enzimas de la ruta de degradación de los hidrocarburos, situando en el caso de los alifáticos a la monooxigenasa en microsomas y en peroxisomas, tanto al Alcohol Graso-

Oxidasa (AGO), como al aldehído deshidrogenasa (Sheller *et al.*, 1998). En el caso de los hidrocarburos aromáticos, solo se menciona el sistema microsomal como sitio de localización de la dioxigenasa. En particular, en las reacciones de degradación de hidrocarburos aromáticos que son catalizadas por oxigenasas, se incluyen cadenas de transferencia de electrones (Gibson y Perales, 2000), eliminación de hidrogeno e inserciones de oxígeno (lo cual puede inducir a la apertura de los anillos aromáticos).

4.7. Técnicas para la Evaluación de Petróleo Crudo en Suelos

4.7.1. Métodos convencionales

Los métodos convencionales para el análisis de hidrocarburos propuestos por la *Environmental Protection Agency* (EPA) (Physical/chemical Methods, 1986), constan habitualmente de dos etapas: extracción y medida. Destacan la extracción Soxhlet (método 3540C), la extracción en fase sólida (método 3535), la extracción con fluidos presurizados (método 3545), la extracción con ultrasonidos (método 3550B) y la extracción con fluidos supercríticos (método 3560).

Los extractantes más usados para la separación de los hidrocarburos son: hexano, metanol, éter etílico, diclorometano, butanol, acetato de etilo, acetona, ciclohexano, o la combinación entre ellos (Zhang *et al.*, 2002; Szolar *et al.*, 2002). La extracción usando Freón 113 (triclorotrifluoretano) fue de las más utilizadas, sin embargo, debido al efecto dañino a la capa de ozono su uso se ha prohibido (Lambert *et al.*, 2001).

Además del uso de disolventes es común utilizar, para tratamientos previos a la extracción, agentes desecantes como el Na_2SO_4 (método 3540C USEPA), o la acidificación de las muestras con H_2SO_4 (Erikson *et al.*, 2001).

A fin de agilizar la separación de los analitos de matrices complejas, como los suelos, existe interés por desarrollar técnicas que permitan una extracción eficiente en un tiempo reducido, un alto nivel de automatización (Camel, 2002) y el empleo de una cantidad pequeña de muestra y disolvente (miniaturización) (Valcarcel y Cárdenas,

2000). Entre las técnicas de separación automáticas destacan la extracción Soxhlet asistida con microondas (Luque-García y Luque, 2001) o automatizada para varias muestras simultáneas (Szolar *et al.*, 2002), extracción con ultrasonidos, extracción con líquidos presurizados, extracción con fluidos supercríticos, extracción con agua subcrítica, o mediante desorción térmica, entre otras (Dea, 1998; Fernández *et al.*, 2002).

Entre las técnicas más utilizados para el análisis de hidrocarburos del petróleo pueden citarse la cromatografía de gases (GC), cromatografía líquida de alta resolución (HPLC), cromatografía de fluidos supercríticos (SFC), cromatografía de capa fina (TLC), cromatografía de gases-espectrometría de infrarrojo con transformada de Fourier (GC-FTIR), espectroscopia de infrarrojo (IR), espectroscopia de ultravioleta (UV) y de fluorescencia, relación de isótopos estables por espectrometría de masas y métodos gravimétrico. De todas las técnicas la CG-MS, LC-MS y GC-FTIR, son las más usadas. Técnicas como la interpretación de fotografía aérea, o de las imágenes de satélite son también consideradas en los estudios ambientales para la identificación de contaminación y sus fuentes, tanto en mar como en tierra (Perrota y Xefteris, 1996).

4.7.2. Métodos modernos

Los métodos modernos presentan, actualmente, un gran interés principalmente por su elevada velocidad de análisis. Algunas de las técnicas basadas en este tipo de generación de señales del conjunto de componentes de una muestra son: espectroscopia en el infrarrojo cercano (NIR) (Wang *et al.*, 1995; Macho y Larrechi, 2002; Zhang *et al.*, 2002; González *et al.*, 2003) e infrarrojo medio (MIR) (Carvalho, 2005), espectroscopia ultravioleta-visible (Thurston *et al.*, 2004) y pirolisis acoplada a espectrometría de masas (Goodacre y Kell, 1996; Radovic *et al.*, 2001, Péres *et al.*, 2002), entre otros. También puede incluirse en este grupo de técnicas la que suele denominarse olfatometría electrónica.

4.7.2.1. Espectroscopia de infrarrojo cercano (NIRS)

En la industria de alimentos, dada la necesidad de conocer el origen y la calidad de los alimentos de consumo humano, se ve necesario realizar la determinación de ciertas propiedades que nos sirvan para caracterizarlos. Los métodos cuantitativos y cualitativos tradicionales de análisis son dificultosos, demandan tiempo, mano de obra y son de un elevado costo (Murray, 1993; Osborne *et al.*, 1998; Cozzolino, 2009). Las propiedades ópticas de los alimentos y en particular la aplicación de la espectroscopia de reflectancia en el infrarrojo cercano (NIRS) se emplea desde la década del 70 en la industria alimenticia, farmacéutica, petroquímica, como alternativa a los métodos químicos y químico-biológicos tradicionales. La NIRS es una técnica rápida, no destructiva, no contaminante, de gran exactitud siempre que se sigan los procedimientos adecuados para crear modelos de clasificación o de predicción (ecuaciones de calibración). En los últimos años se han desarrollado numerosas aplicaciones NIRS para evaluar la composición el procesamiento y certificación de la calidad de los alimentos. El método utiliza la región del infrarrojo cercano de longitudes de onda entre los 700 y 2500 nanómetros (nm) del espectro electromagnético. Este método se basa en que, cuando la luz (energía) incide sobre una muestra, una parte de los fotones se transmite a través de la misma, siendo el resto absorbida. La absorción de energía por la muestra produce que los enlaces entre C–H, N–H, O–H, S–H, C=O y C=C, principales constituyentes de la estructura básica de las sustancias orgánicas, vibren en distintas formas. La interacción de la energía con la materia obedece a la ley de Lambert-Beer, que establece que la absorbancia a cualquier longitud de onda es proporcional al número o concentración de moléculas absorbentes presentes en el camino recorrido por la radiación (Cozzolino, 2002; Osborne, *et al.*, 1998).

Esto determina que para un material de naturaleza química heterogénea el espectro obtenido en la región del infrarrojo cercano sea una combinación de bandas de absorciones parciales sobrepuestas o muy cercanas, que suelen confundirse en una línea suavizada, en que se encuentran picos, valles y curvaturas en forma de hombro (Alomar y Fuschlocher, 1998). Al desarrollar una calibración NIRS, se relaciona

mediante un algoritmo la información espectral (óptica) con la información de la composición físico-química (método de referencia) a través de la aplicación de modelos estadísticos como son la regresión múltiple, los componentes principales y los mínimos cuadrados parciales.

En el área ambiental se evalúa la factibilidad de usar la espectroscopia NIR en la predicción de contaminantes en suelos agrícolas, ya que los métodos convencionales para investigación de la contaminación de suelos son muy laboriosos y caros. Wu *et al.* (2007) determinaron una predicción segura para los metales Ni, Cr y Co, moderada para Cu, As, Zn y Pb y pobre para Cd, a través del uso del NIR.

La espectroscopia de infrarrojo cercano (NIR) es usada en muchas áreas, los espectros NIR de una muestra son medidos típicamente por escáneres en cientos de longitudes de onda espaciados igualmente. La información en la curva es usada para predecir la composición química de la muestra por extracción de información relevante de muchos picos sobrepuestos. Aproximaciones estándar, tales como el Análisis Lineal Discriminante (LDA) han sido descritos, sin embargo, estos métodos fallan con muchas variables, así que necesitan tomarse diferentes aproximaciones. Una solución es reducir la dimensión de la matriz de predicción y entonces aplicar LDA. Diversos métodos de clasificación han sido comparados usando espectros de masas, incluyendo análisis discriminante lineal y cuadrático y la clasificación por el método del árbol. Autores como Nikolich *et al.* (2001), enfatizan la necesidad de métodos para eliminar el ruido de los datos y seleccionar las características relevantes. La transformación de pequeñas ondas y el análisis discriminante estándar son aplicadas para remover el ruido.

El espectro NIR muestra sobretonos y combinación de las bandas de estos grupos. Desafortunadamente, las bandas de absorción en la región NIR son amplias y se superponen, lo cual significa que las técnicas convencionales de calibración univariadas, usando solo una longitud de onda por componente para evaluaciones, no pueden ser aplicadas en casos de superposición de bandas. El desarrollo de herramientas estadísticas más sofisticadas, como el método multivariado de regresión

de mínimos cuadrados parciales (PLS), da la posibilidad para la amplia aplicación de espectroscopia NIR a muchos laboratorios (Nikolich *et al.*, 2001).

La espectroscopia NIR ofrece un gran número de ventajas para análisis cualitativos y cuantitativos y aplicaciones de control de procesos. Entre las principales ventajas se tiene: 1) no preparar la muestra para el análisis, 2) reducción del error analítico, 3) no hay residuos, 4) costos reducidos, 5) mediciones y análisis más rápidos, 6) componentes ópticos no giroscópicos, 7) mediciones remotas por fibra óptica, 8) mayor exactitud en los análisis y 9) fácil uso (Batten, 1998; Blanco y Villarroya, 2002).

Con la técnica de reflectancia de infrarrojo cercano, la muestra por analizar está sujeta a radiación NIR, la superficie de la muestra absorbe parte de la energía, mientras que el resto es dispersado en todas direcciones. La luz dispersada da información sobre la composición de la muestra en la prueba y el espectro resultante es usado en el análisis cualitativo y cuantitativo. Los mejores intervalos de frecuencia están en la región del espectro en el cual la intensidad de luz reflejada se correlaciona mejor con las concentraciones de las especies químicas analizadas (Alomar y Fuchslocher, 1998).

4.8. Quimiometría

La Quimiometría puede definirse como una disciplina química que utiliza métodos matemáticos y estadísticos para diseñar procedimientos de medida y para obtener la mayor cantidad de información química posible a partir del análisis de datos (Massart *et al.*, 1997). Trata, específicamente, de todos aquellos procesos que transforman señales analíticas y datos más o menos complejos en información. Utiliza métodos de origen matemático, estadístico y otros procedentes del campo de la lógica formal para conseguir sus fines. Por todo ello, la quimiometría se sitúa en un campo interdisciplinario. Aunque sus métodos y herramientas provienen de otras disciplinas (como, de hecho, ocurre habitualmente en la química analítica), claramente los fines de la quimiometría están ligados a la química y su éxito depende de los problemas químicos que sea capaz de resolver (Miller, 2005).

Dentro de la Quimiometría pueden citarse los métodos de diseño experimental y exploración de datos, los métodos de reconocimiento de pautas, calibración y regresión simple y multivariante así como los sistemas expertos e inteligencia artificial, entre otros.

En análisis multivariante cada muestra puede considerarse como un punto en un espacio multidimensional. El objetivo es extraer la información útil de todo el conjunto de las variables medidas en una experiencia. El diseño de experimentos es un complemento de este análisis porque proporciona una estrategia para diseñar la experimentación de un proceso. Esto permite obtener la mayor información útil posible con el menor número de experimentos. Además, la calidad de los parámetros calculados depende exclusivamente del diseño seleccionado y ésta es óptima (menor varianza) cuando la matriz del diseño es ortogonal, es decir, las concentraciones de los diferentes compuestos no están correlacionadas (Massar *et al.*, 1997).

Los métodos de reconocimiento de pautas tienen la finalidad de clasificar o agrupar muestras según las características de las mismas. De acuerdo con Brereton (1990), pueden distinguirse dos tipos de técnicas de reconocimiento de pautas: supervisadas y no supervisadas.

Las técnicas no supervisadas tienen como objetivo determinar relaciones entre muestras sin utilizar ninguna información previa con respecto a la pertenencia a determinadas clases. Permiten conocer la existencia de estructuras en los datos y tener una idea de la utilidad de la información contenida en los mismos. Entre estos métodos se encuentran el análisis de agrupamientos, una de cuyas modalidades es el agrupamiento jerárquico (HCA) en el que las muestras se asocian en función de unos índices de similitud, y el análisis de componentes principales (PCA), en el que se reduce la dimensión de un conjunto de datos mediante el cálculo de combinaciones lineales de las variables originales (Brereton, 2002).

Las técnicas de reconocimiento de pautas supervisadas tienen como fin desarrollar

reglas de clasificación para muestras (objetos) desconocidas a partir de un grupo de muestras conocidas que están caracterizadas por los valores de las variables medidas (Hopke, 2003). El éxito de la predicción depende de que las reglas de clasificación y los límites marcados por ellas sean adecuados. El objetivo final es predecir a qué grupo (clase) pertenece una muestra desconocida. Algunas de las técnicas más importantes son el análisis discriminante lineal (LDA), modelado independiente de clases (SIMCA), “support vector machines” (SVM), el método de los k vecinos más próximos (k-NN) y las redes neuronales artificiales (ANN), entre otros.

4.8.1. Reconocimiento de pautas no supervisadas

4.8.1.1. Análisis de agrupamientos

En esta técnica de análisis exploratorio se calcula y compara la similitud entre las muestras. Algunas de las posibilidades para determinar la similitud entre los objetos se describen a continuación (Brereton, 2003).

- Coeficientes de correlación entre las muestras.

Un coeficiente de correlación de 1 implica que las muestras tienen características idénticas. Este coeficiente se obtiene cuando una muestra se compara con ella misma. Los valores pueden oscilar entre -1 y 1 aunque es habitual trabajar con el cuadrado de este número, es decir, el cuadrado del coeficiente de correlación.

- Distancia Euclídea.

La distancia Euclídea entre dos muestras k / l está definida por:

$$d_{kl} = \sqrt{\sum_{j=1}^J (x_{kj} - x_{lj})^2} \quad (1)$$

donde J representa el número total de variables medidas y x_{kj} y x_{lj} , son los valores de la variable j para las muestras k y l , respectivamente. Esta medida es en cierta forma, contraria al coeficiente de correlación y estrictamente es una medida de “no similitud”. No hay restricciones para su valor si es siempre positivo. La ecuación (2) es la notación matricial de la ecuación (1):

$$d_{kl} = \sqrt{(\mathbf{x}_k - \mathbf{x}_l) \cdot (\mathbf{x}_k - \mathbf{x}_l)'} \quad (2)$$

- Distancia de Manhattan.

La distancia entre dos muestras k y l se determina según:

$$d_{kl} = \sum_{j=1}^J |x_{kj} - x_{lj}| \quad (3)$$

Las diferencias entre las distancias Euclídea y de Manhattan se ilustran en la Figura 6 en el caso de dos variables. Esta distancia es siempre mayor o al menos igual que la Euclídea.

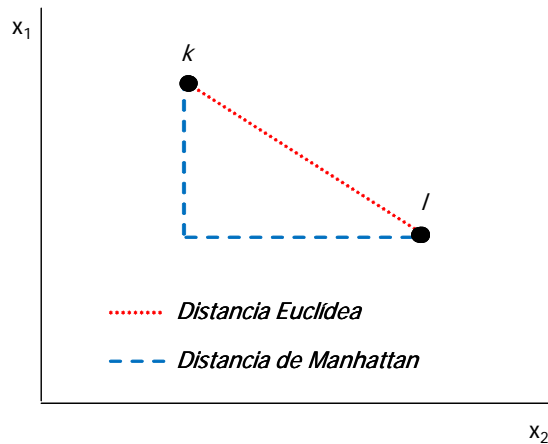


Figura 6. Distancia Euclídea y de Manhattan para las muestras k y l .

- Distancia de Mahalanobis.

Se asemeja a la distancia Euclídea pero tiene en cuenta que algunas variables pueden estar correlacionadas y por lo tanto miden en cierto grado la misma propiedad. La distancia entre dos objetos k y l se puede definir como:

$$d_{kl} = \sqrt{(x_k - \bar{x}_l)C^{-1}(x_k - \bar{x}_l)'} \quad (4)$$

donde C es la matriz varianza-covarianza de las variables. Es una matriz simétrica cuyos elementos representan las covarianzas entre pares de variables. Los elementos de la diagonal corresponden a la varianza de cada variable. La matriz es de dimensiones $J \times J$.

Este método no es fácilmente aplicable cuando el número de variables excede al de muestras analizadas debido a que la matriz C no tiene inversa. Una posible solución requeriría seleccionar un número más pequeño de variables. Su utilización requiere trabajar con componentes principales en lugar de con las variables originales (ver próxima sección, PCA).

El análisis de agrupamientos puede dividirse en dos tipos de técnicas: agrupamientos jerárquicos y no jerárquicos.

En los agrupamientos jerárquicos se generan dendrogramas en los que se visualizan las distancias entre muestras y en los que pueden considerarse diferentes grupos en función de determinados niveles de similitud. En los agrupamientos no jerárquicos es necesario prefijar el número de grupos en los que se quieren dividir las muestras y éstos se generan mediante un proceso iterativo en el que la pertenencia de cada muestra a un grupo concreto se va modificando hasta llegar a una situación estable.

En el caso del análisis de agrupamientos jerarquizados (HCA), después de calcular las distancias entre todos los pares de muestras, el siguiente paso es generar los agrupamientos. Los pasos que se deben seguir son los siguientes (Brereton, 2003).

- a) Encontrar las dos muestras más similares, esto es, aquellas entre las que existe la menor distancia.
- b) Formar un grupo con estas muestras. Ahora se sustituye el grupo ya formado por un punto que lo represente. Para seleccionar al representante se suelen aplicar

los métodos (Ramis *et al.*, 2001) que se ilustran en la Figura 7. En el método del centroide, los puntos que ya forman grupo quedan representados por su centroide. Sin embargo en los otros métodos, el grupo queda representado por aquel punto que presente la menor distancia con respecto a la nueva muestra en el caso de la unión simple y por el de mayor distancia si se aplica el método de la unión completa. A continuación se recalculan las nuevas distancias entre el nuevo grupo formado y todas las demás muestras buscando de nuevo el valor más pequeño.

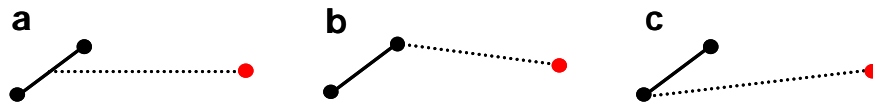


Figura 7. Unión de un grupo con una muestra aislada de acuerdo con los métodos del centroide (a), de la unión simple (b) y de la unión completa (c).

- c) Los siguientes pasos consisten en agrupar los datos de manera similar hasta que todos los objetos se unan.

El resultado es una representación gráfica llamada dendrograma. Es una estructura en forma de árbol donde la longitud de las ramas es proporcional a la distancia entre los grupos unidos. La forma del dendrograma depende del método de unión. Habitualmente, las distancias entre muestras se transforman según la siguiente escala de similitud:

$$S = 1 - \frac{d_{kl}}{d_{max}} \quad (5)$$

donde d_{max} es la mayor distancia del conjunto de datos.

En el ejemplo de la Figura 8 los objetos se dividen formando fundamentalmente 2 agrupamientos (A y B). Las muestras 1, 4, 7 y 2 presentan características similares entre sí pero distintas a las del otro grupo, constituido por las muestras 6, 8,5 y 3.

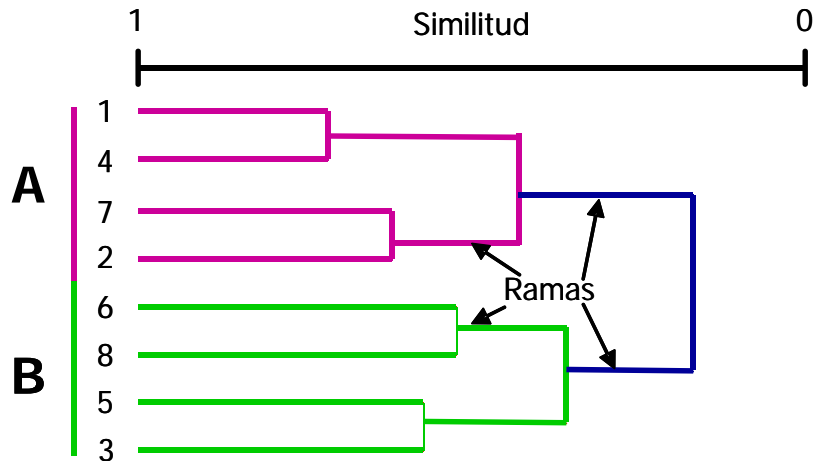


Figura 8. Dendrograma que representa la cercanía entre muestras.

4.8.1.2 Análisis de Componentes Principales (PCA)

Es una técnica de reconocimiento de pautas no supervisadas que permite encontrar la estructura “latente” en una tabla de datos cuya lectura directa lo impide. El método trata de reducir la dimensión del conjunto de datos construyendo unas nuevas variables, ortogonales entre sí, que reciben el nombre de componentes principales (PCs) y que son combinaciones lineales de las variables primitivas.

En este análisis la matriz de datos X formada por I objetos o muestras y J variables (dimensión $I \times J$) se descompone (Massart *et al.*, 1997) en el producto de las matrices T y P que contienen la información relacionada con las muestras y con las variables respectivamente.

$$X = T \cdot P + E \quad (6)$$

En la expresión (6) la matriz T contiene las puntuaciones, valores (*scores*) de cada muestra en las nuevas variables, de I objetos en A componentes principales. La matriz P contiene los pesos, coeficientes (*loadings*) en las variables originales para generar las nuevas, de J variables sobre A componentes principales. E es una matriz residual. Las matrices de puntuaciones y pesos consisten en una serie de vectores columna y fila

respectivamente según muestra la Figura 9. Ambas matrices están formadas por tantos vectores como PCs puedan calcularse. Estos vectores se representan como t_a y p_a donde a es el componente considerado y reciben el nombre de vectores principales. Si la matriz de datos originales es de dimensiones $I \times J$, el número máximo de PCs que pueden ser calculados es el valor de la dimensión más pequeña.

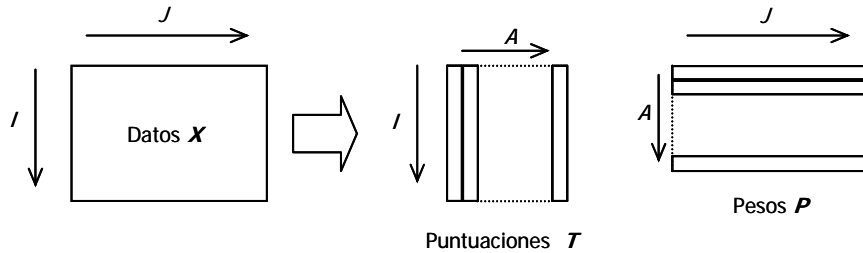


Figura 9. Análisis de componentes principales (PCA).

El objetivo de la técnica es que el número de componentes principales elegido explique la parte determinante de la matriz de datos de modo que los residuales sean lo más pequeños posibles. En el caso de estudiar una mezcla de compuestos (Brereton, 2003), el número de PCs debería ser idéntico al número de compuestos existentes. Sin embargo esto no es generalmente tan simple y debemos tener en cuenta otros factores como por ejemplo el ruido que pueden presentar los datos. Los primeros componentes son sin duda los más importantes. El primero explica el mayor porcentaje de varianza de los datos, el segundo el mayor de la restante y así sucesivamente. Mediante el proceso de selección del número de PCs se retiene la información significativa de las matrices y se separa de aquella relacionada con el ruido.

La importancia de cada componente puede medirse mediante la suma de los cuadrados de las puntuaciones según la expresión:

$$g_a = \sum_{i=1}^I t_{ia}^2 \quad (7)$$

donde g_a es el valor propio (*eigenvalue*) del componente principal a . Además la suma de todos los valores de g_a es la suma de los cuadrados de los valores de la matriz de datos X :

$$\sum_{a=1}^K g_a = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 \quad (8)$$

donde K corresponde al valor más pequeño de I ó J . Frecuentemente estos valores se presentan como porcentajes tal y como muestra la expresión (9).

$$V_a = 100 \cdot \frac{g_a}{\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2} \quad (9)$$

donde V_a es el porcentaje de información modelada por el componente a .

El porcentaje acumulado es utilizado a menudo para determinar la proporción de información que ha sido modelada.

Tras aplicar esta técnica, los datos originales quedan matemáticamente modelados por medio de unos pocos componentes principales. Usando un número A de componentes es posible establecer un modelo para cada elemento de X de la forma:

$$x_{ij} = \sum t_{ia} \cdot p_{aj} + e_{ij} = \hat{x}_{ij} + e_{ij} \quad (10)$$

que es la expresión no matricial de la ecuación (6).

Una forma más clara de visualizar esta sección es buscar una interpretación gráfico-geométrica (Blanco y Cerda, 1988) a este análisis. Así, un objeto queda definido por un punto en el hiperespacio constituido por J variables. Esta técnica trata de reducir la dimensión construyendo unas nuevas variables de forma que los objetos estén situados en el nuevo espacio, de forma similar a como lo estaban en el original. Esto produce un giro en el hiperespacio definido por las J variables de forma que la primera variable nueva (primer eje calculado) contenga la máxima información, la segunda (ortogonal a la primera) el máximo de la restante y así sucesivamente. La matriz de puntuaciones está constituida por las proyecciones de cada uno de los objetos del conjunto de datos en los nuevos ejes o componentes principales. La representación de las puntuaciones de los objetos sobre los componentes principales permite reconocer agrupamientos de

las muestras. Esto se ilustra en la Figura 10.

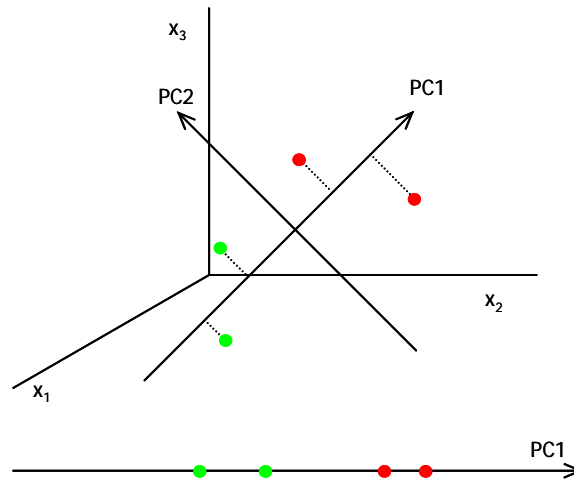


Figura 10. Representación de las puntuaciones de dos clases de objetos en los dos primeros componentes principales

4.8.2 Reconocimiento de pautas supervisadas

Los métodos de reconocimiento de pautas supervisadas son también conocidos como métodos de clasificación. Su objetivo es clasificar muestras desconocidas en determinados grupos o clases a partir de la información suministrada por un grupo de entrenamiento. Aunque hay numerosos algoritmos es habitual utilizar una estrategia común para clasificar las muestras independientemente de la herramienta matemática que se utilice. La estrategia consta de tres etapas (Brereton, 2003).

- a) Crear un modelo con el grupo de entrenamiento.

El primer paso es generar un modelo matemático que relacione los valores de las variables medidas del conjunto de muestras con el grupo al que pertenecen. Estos objetos reciben el nombre de conjunto de entrenamiento.

- b) Comprobar la capacidad predictiva del modelo.

Después de generar el modelo se determina su capacidad predictiva. Un parámetro representativo puede ser el porcentaje de objetos correctamente clasificados.

Esto puede realizarse siguiendo distintos criterios. Dos de los más habituales se describen a continuación:

- Validación cruzada (*cross validation*).

Es una forma de validación interna del modelo. Una muestra o un conjunto pequeño de ellas se excluyen del conjunto de entrenamiento. Ahora se construye un nuevo modelo sin las muestras excluidas y se comprueba la capacidad de predicción del modelo sobre ellas. El procedimiento se repite hasta que todas las muestras han quedado fuera del modelo al menos una vez.

- Validación externa (*test set*).

Para comprobar la calidad del modelo se realiza la predicción de un grupo independiente de muestras que no ha sido utilizado en ninguna etapa de la construcción del modelo. Se asume que estas muestras son desconocidas y se comparan los valores obtenidos tras aplicar el modelo con los valores reales conocidos.

Cuando el modelo generado no es satisfactorio existen diferentes posibilidades de mejorarlo. Por un lado se puede utilizar otro algoritmo de cálculo o realizar una selección de variables ya que algunas pueden estar ampliamente influidas por el ruido. En lugar de utilizar todas las variables medidas se seleccionan sólo aquellas que aportan mayor información. El resto podrían estar relacionadas con el ruido.

- c) Predicción del conjunto desconocido de muestras.

Una vez que se dispone de un modelo satisfactorio, éste podrá ser aplicado a muestras desconocidas. Un aspecto importante que se debe tener en cuenta es la estabilidad del modelo con el tiempo, es decir, cuando el conjunto desconocido de muestras se analiza semanas o incluso meses después que las correspondientes del grupo de clasificación. En algunas ocasiones es posible que existan variaciones en la respuesta instrumental registrada debido a problemas de inestabilidad de la señal con el transcurso de los días.

Esto puede tener seria influencia sobre la capacidad del modelo para clasificar satisfactoriamente una muestra. En problemas de clasificación es habitual realizar algún tipo de normalización de los datos antes de su análisis.

4.8.2.1. Análisis discriminante lineal (LDA)

La clasificación univariante es el método más sencillo. Solo se dispondría de una variable para dividir los objetos en los grupos. Sin embargo, a menudo se requiere más de una medida para determinar a qué grupo pertenece una muestra. En la Figura 6a los objetos representados en rojo podrían considerarse distintos de los representados en verde. Ninguna de las variables (x_1 , x_2) por separado es capaz de discriminar los objetos. Ambas son esenciales. En la Figura 11 es posible dibujar una línea (línea discriminante) entre ambos grupos. Los objetos por encima de la línea pertenecen a la clase A y los que están por debajo a la clase B. Gráficamente esto se puede representar mediante las proyecciones de los objetos sobre la línea perpendicular a la discriminante tal y como se muestra en las Figuras 8b y 8c. Ahora es fácil calcular el centro de cada clase y la distancia de una muestra desconocida a ambos centroides. La muestra desconocida se asignará al grupo más cercano.

Los resultados obtenidos con este análisis deben ser cuidadosamente estudiados ya que aquellas muestras con características de ambas clases siempre serán asignadas, exclusivamente, a una de ellas e incluso, las muestras que no pertenecen a ninguna clase siempre serán asignadas a una de ellas (Coomans *et al.*, 1979).

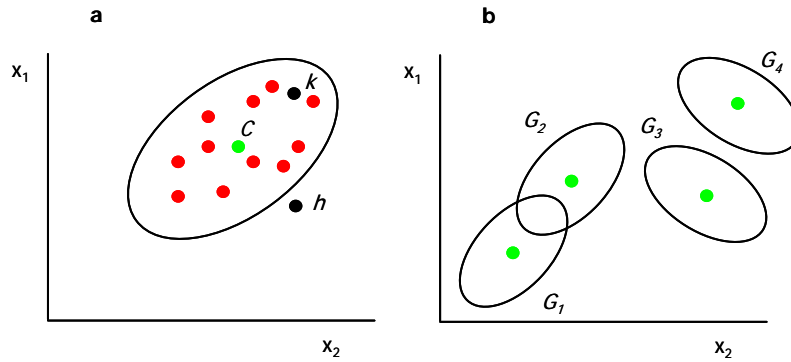


Figura 11. (a) Dos clases de objetos definidos por las variables x_1 y x_2 . (b) Línea discriminante que separa las clases de objetos. (c) Proyecciones de las muestras en la recta perpendicular a la línea discriminante y cálculo de centroides.

En muchas ocasiones se utiliza la distancia de Mahalanobis para medir distancias entre grupos de muestras. Las Figuras 12a y 12b ilustran las principales características de esta distancia. La distancia Euclídea entre el centro del grupo (punto C) y los objetos k y h de la Figura 8a es idéntica pero k forma parte del grupo mientras que h no. Sería deseable que la distancia de C a h fuera mayor que la distancia entre C y k . La distancia de Mahalanobis indica que el punto k está más “cerca” que h de C porque está situado en la dirección del eje mayor de la elipse mientras que h no. Los objetos situados dentro de la elipse tienen valores de x_1 y x_2 que están fuertemente correlacionados. Este no es el caso de h . De manera similar en la Figura 11b, los grupos G_1 y G_2 están “más cerca” entre ellos que G_3 y G_4 según la distancia de Mahalanobis aunque presenten la misma distancia Euclídea. Todos los grupos tienen la misma forma y el mismo volumen pero G_1 y G_2 se solapan mientras que G_3 y G_4 no.

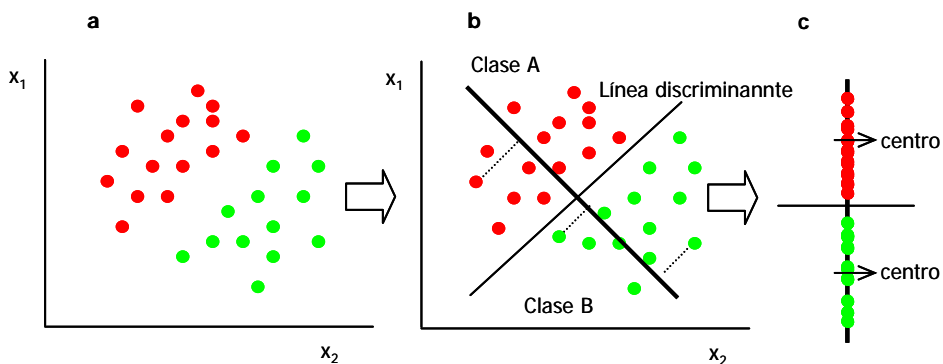


Figura 12. Distancia de Mahalanobis que determina la similitud entre dos variables aleatorias multidimensionales.

La distancia de Mahalanobis de una muestra k al centroide de una clase (A) viene definida por la expresión:

$$d_{kA} = \sqrt{(x_k - \bar{x}_A)C_A^{-1}(x_k - \bar{x}_A)^T} \quad (11)$$

donde x_k es el vector fila de la muestra k que recoge el valor de todas las variables y \bar{x}_A el que corresponde al centroide de esta clase. C_A es la matriz varianza-covarianza para este grupo.

En el análisis discriminante lineal pueden considerarse dos etapas: construcción de las funciones discriminantes y determinación de las reglas de clasificación. Esta técnica establece unas funciones discriminantes que forman un espacio reducido con respecto a las variables originales. La dimensión de este espacio es, generalmente, una menos que el número de clases que posea el grupo de entrenamiento. Para la construcción de las funciones se considera el caso donde existan dos clases de objetos A y B caracterizados por una serie de variables. El objetivo es ponderar y combinar linealmente las variables de manera que los conjuntos de muestras se diferencien todo lo posible reduciendo a la vez el número de dimensiones con una pérdida mínima de información. Es posible calcular distintas funciones discriminantes lineales para separar dos clases de objetos aunque, en este caso sencillo, es suficiente una única función del tipo:

$$D = \sum_{j=1}^J w_j y_j \quad (12)$$

donde D es la función discriminante, w_j es el coeficiente, y_j es la variable y J es el número de variables.

La obtención de los coeficientes resulta fundamental ya que no todos los ejes proporcionan una diferencia adecuada entre clases. Esto se consigue seleccionando de entre todas las posibles funciones aquellas que producen las mayores distancias de Mahalanobis entre grupos diferentes.

Si el número de clases es superior a dos se necesita un mayor número de funciones discriminantes. Éstas deben de ser ortogonales entre sí, de forma que no compartan información.

Una vez determinadas las funciones, se necesita desarrollar una regla matemática de clasificación que divida el espacio en tantas regiones como clases. Una posible regla, mencionada anteriormente, puede ser considerar la distancia de una muestra a los centros de las clases de forma que la muestra desconocida se asigna a la clase cuyo centroide esté más cercano.

4.8.2.2 Modelo suave independiente de analogía de clases (SIMCA)

Uno de los métodos de reconocimiento de pautas supervisadas más utilizados es SIMCA (Massart *et al.*, 1997; Brereton, 1990).

Mientras que en otros métodos de clasificación cada muestra se asigna solamente a una clase y ninguna queda fuera de alguna de ellas, éste permite que el resultado de la clasificación para una muestra sea que no pertenece a ninguna de las clases o que pertenece a más de una (con la misma o diferente probabilidad). En muchas situaciones reales no es posible asignar una muestra concreta a alguna de las clases o existe la posibilidad de que algunas muestras pertenezcan a dos clases simultáneamente.

El método se basa en construir un modelo independiente para cada una de las clases mediante PCA. El número de componentes principales utilizados para cada clase puede fijarse si se dispone de información previa o mantener un número de ellos que explique un determinado porcentaje de la varianza de los datos. En la Figura 13 se presentan dos clases de objetos, A y B. Al aplicar PCA a la clase A (puntos verdes) se obtiene un modelo de dos componentes (un plano), mientras que la clase B (puntos rojos) se ajusta a un modelo de una única variable (una recta). Las regiones de confianza aparecen alrededor de cada clase.

La asignación de los nuevos objetos se hace comprobando si están localizados en la

región de confianza de algún modelo. Así el objeto k de la Figura 10 se puede asignar a la clase B mientras que h no pertenece a ninguna clase, es un “outlier”.

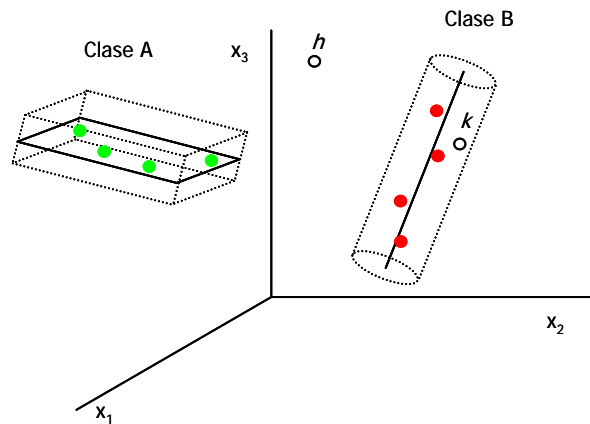


Figura 13. Modelos de dos clases de objetos A y B en el espacio de sus componentes principales, regiones de confianza y predicción de dos muestras.

Para cada clase y su modelo correspondiente se define la sensibilidad como el porcentaje de muestras que perteneciendo a esa clase son reconocidas correctamente por el modelo matemático. Análogamente se define la especificidad como el porcentaje de muestras que perteneciendo a otra clase son reconocidas como ajenas al modelo.

Los diagramas de Coomans permiten visualizar estos índices ya que son representaciones bidimensionales de las distancias de las muestras a los modelos. La Figura 14 es un esquema de estos diagramas. En el eje de abscisas se representa la distancia a la clase A y en el de ordenadas a la clase B. La recta horizontal discontinua delimita el recinto de la clase B mientras que la línea vertical discontinua delimita el recinto de la clase A. El cuadrado inferior izquierdo (1) es el recinto común a ambas clases y cuantas menos muestras tengan sus coordenadas en él, mayor será la especificidad de los modelos. El cuadrado superior derecho (2) representa el espacio que queda fuera de ambos modelos. Las muestras con coordenadas en él son muestras extrañas a los modelos generados.

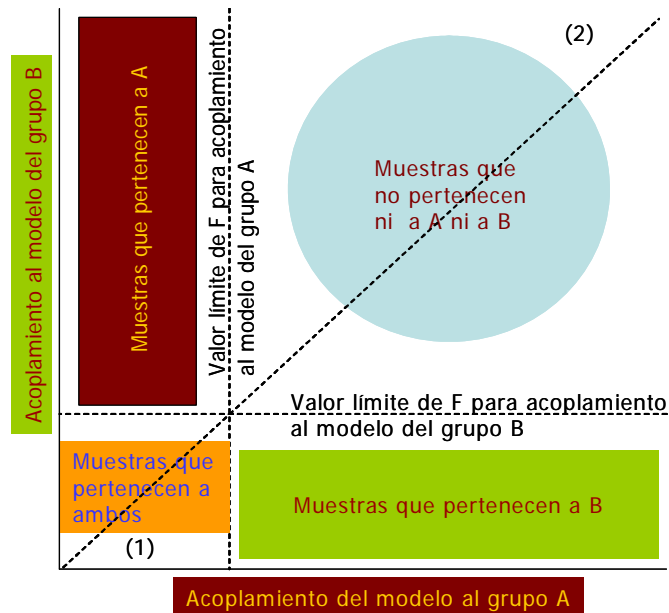


Figura 14. Diagrama de Coomans.

Cuando SIMCA se utiliza únicamente como técnica de clasificación, todas las muestras que se encuentren por encima de la diagonal serán asignadas a la clase A y las muestras por debajo a la clase B.

4.9 Modelos de Calibración

4.9.1 Calibración univariante

La calibración univariante (Massart *et al.*, 1997; Brereton, 2000; Miller y Miller, 2000) permite obtener modelos capaces de predecir una variable respuesta conociendo los valores de otra correlacionada con ella. La expresión característica (Brereton, 2000) de este tipo de calibración es:

$$X = c's \tag{13}$$

Donde X es un vector que está formado por las respuestas analíticas para el número de muestras analizadas y c es el vector que corresponde a las concentraciones. Ambos vectores tienen la misma longitud, I , que es igual al número de muestras el escalar S

relaciona ambos parámetros.

En muchas situaciones será apropiado incluir en término adicional en el modelo de calibración, la ordenada en el origen.

La calidad de las predicciones puede determinarse mediante la expresión:

$$E = \sqrt{\frac{\sum_{i=1}^I (c_i - \hat{c}_i)^2}{d}} \quad (14)$$

donde $(c_i - \hat{c}_i)$ es la diferencia entre las concentraciones añadida y predicha para la muestra i . "a" son los grados de libertad. En el caso de la calibración invariante el número de los grados de libertad es igual al número de experimentos realizados menos el número de parámetros calculados en el modelo. Este error absoluto puede representarse como porcentaje mediante la expresión:

$$E\% = \frac{E}{C} \cdot 100 \quad (15)$$

Donde \bar{c} es media de las concentraciones de todas las muestras.

A partir de los modelos clásicos que incluyen ordenada en el origen es posible determinar tres varianzas de interés (Ramis, 2001).

- La varianza explicada por el modelo.

Es la suma de cuadrados de (SC) de las desviaciones de los valores predichos respecto a las coordenadas del centroide dividida por el número de grados de libertas que es uno.

$$S_{\text{exp}}^2 = \sum_{i=1}^I (\hat{x}_i - \bar{x}_i)^2 = sc_{\text{exp}} \quad (16)$$

➤ La varianza residual.

Es la suma de cuadrados de las desviaciones de los valores reales respecto a los predichos divididos entre el número de grados de libertad, que en este caso es el número de experimentos realizados menos dos. Este tipo de varianza es una medida de la dispersión de las medidas a lo largo de la línea de regresión ajustada y representa la varianza en la respuesta que el modelo no tiene en cuenta.

$$S_{tot}^2 = \frac{\sum_{i=1}^I (x_i - \bar{x}_i)^2}{I-1} = \frac{SC_{tot}}{I-1} \quad (17)$$

➤ La varianza total.

La varianza total de los puntos en ausencia de modelo se calcula mediante la suma de cuadrados de las desviaciones de los puntos respecto al centroide dividida por el número de grados de libertad, $I - 1$.

$$S_{tot}^2 = \frac{\sum_{i=1}^I (x_i - \bar{x}_i)^2}{I-1} = \frac{SC_{tot}}{I-1} \quad (18)$$

La relación (Brereton, 2003) entre las tres sumas de cuadrados es:

$$SC_{tot} = SC_{exp} + SC_{res} \quad (19)$$

Lo que manifiesta que el modelo divide la varianza total en dos, la explicada por él y la residual. Fácilmente se puede demostrar (Ramis, 2001) que la relación entre SC_{exp} y SC_{tot} equivale al cuadrado del coeficiente de correlación lineal:

$$r^2 = \frac{\sum (x_i - \bar{x}_i)^2}{\sum (x_i - \bar{x}_i)^2} = \frac{SC_{exp}}{SC_{tot}} \quad (20)$$

El parámetro r^2 se denomina coeficiente de determinación y expresa el porcentaje de

variación total de los datos que es explicado por el modelo. Se utiliza para comprobar el porcentaje de información modelada por la regresión pero necesariamente no implica que el modelo más adecuado para ajustar los datos sea el lineal incluso cuando sus valores próximos a 1. Además de obtener un buen coeficiente de correlación es importante comprobar que el modelo generado no contiene fallo de ajuste. Para ello es necesario tener replicas de cada nivel de concentración.

La secuencia que se representa a continuación corresponde a un conjunto de muestras con n niveles de concentración diferentes, repetidos cada uno p veces. X_{ij} son los valores experimentales obtenidos para las muestras.

$$\begin{array}{l}
 C_1 \quad X_{11} \dots X_{1j} \dots X_{1p} \\
 C_2 \quad X_{21} \dots X_{2j} \dots X_{2p} \\
 \cdot \\
 \cdot \\
 \cdot \\
 C_i \quad X_{i1} \dots X_{ij} \dots X_{ip} \\
 \cdot \\
 \cdot \\
 \cdot \\
 C_n \quad X_{n1} \dots X_{nj} \dots X_{np}
 \end{array}$$

Ahora es posible dividir la variación residual, SC_{res} , en dos partes:

- Variación debido al error puro.

Mide la variación debida al error experimental según la expresión:

$$SC_{error} = \sum_{ij} (x_{ij} - \bar{x}_i)^2 \quad (21)$$

Donde \bar{x}_i es la media de las medidas realizadas para cada nivel de concentración.

- Variación debida al cambio de ajuste

Es un componente que mide la variación de la media de las muestras repetidas para cada nivel de concentración a lo largo de la línea de regresión mediante la expresión.

$$SC_{fall.aju} = SC_{res} - SC_{error} \quad (22)$$

Las varianzas asociadas a estos factores se obtienen dividiendo las sumas de cuadrados entre el número de grados de libertad tal y como se indica en las siguientes expresiones:

$$S_{error} = \frac{SC_{error}}{n(p-1)} \quad (23)$$

$$S^2_{fall.ajus} = \frac{SC_{res} - SC_{error}}{n-2} \quad (24)$$

El posible fallo de ajuste se puede comprobar comparando la relación $F = \frac{S^2_{fall.ajus}}{S_{error}^2}$

obtenida experimentalmente con el valor crítico tabulado de F a (n-2) y n (p-1) grados de libertad y un determinado nivel de significación. Si la diferencia es significativa (la varianza del fallo de ajuste es mayor que la varianza debida al error puro) se puede concluir que el modelo es inadecuado debido a que la variación de las medias correspondientes a las muestras repetidas con el mismo nivel de concentración no puede ser explicado en términos de la incertidumbre del error puro. En este caso es aconsejable adoptar un modelo diferente. En caso contrario el modelo no representa fallo de ajustes y es adecuado para representar el conjunto de muestras estudiadas.

4.9.2. Regresión lineal múltiple

La regresión lineal múltiple (MLR) es una extensión del modelo anterior la que se utiliza más de una variable predictora o respuesta instrumental. Las principales razones para utilizar estos nuevos métodos son dos. En primer lugar es posible que exista más de un compuesto en las muestras que se quieran analizar. En estas circunstancias es

necesario utilizar más de una respuesta.

Para n compuestos al menos deben ser medidas n respuestas diferentes a no ser que se conozca que la concentración de algunos de ellos está correlacionada. La segunda razón está relacionada con el hecho de que cada respuesta contiene algún tipo de información y, aunque algunas pueden contener ruido o interferentes desconocidos, los modelos que incluyen más de una variable predictora proporcionan, generalmente mejores resultados que aquellos contruidos con una sola.

Esta técnica proporciona buenos resultados cuando se incluyen en el modelo todos los compuestos que contribuyen a la señal medida. En caso contrario las predicciones empeoran notablemente. Supongamos una mezcla de diez compuestos. La regresión lineal múltiple proporciona resultados adecuados si se dispone de información suficiente de los diez compuestos. Sin embargo si solo se sabe de la presencia de tres de ellos, las predicciones empeorarán ya que el resto de los compuestos distorsionan el modelo. Los compuestos no modelados contribuyen al cálculo de los coeficientes de las variables predictoras aunque luego se afirma que el modelo formado corresponde a un número menor de compuestos. Este problema podría solucionarse (Brereton, 2000), en parte, si se seleccionan exclusivamente las variables más importantes de los tres compuestos que se estudian. La necesidad de modelar todos los compuestos significativos es la mayor limitación de esta regresión.

4.9.3. Regresión por mínimos cuadrados parciales (PLS)

La regresión por mínimos cuadrados parciales es a menudo considerada como la técnica de regresión multivariante (Martens, 1996) más importante y a diferencia de la técnica anterior no requiere un modelo exacto para todos los compuestos.

El método está basado en análisis de componentes principales y no requiere de detalles de todos los compuestos presentes en las muestras aunque será necesario realizar una estimación del número de componentes principales significativos que no se

corresponde necesariamente con el número de especies químicas (Brereton, 2000).

Uno de los aspectos más importantes de esta técnica de calibración frente a otras como la regresión por componentes principales (PCR) es que tiene en cuenta no solo los posibles errores de bloque de datos X , sino también los errores relacionados con la estimación de la concentración.

Existen fundamentalmente dos métodos que se denominan PLS1 y PLS2. A continuación se describen ambos.

4.9.2.1. PLS1

Los bloques de matrices X y c se descomponen en sus componentes principales de acuerdo con las siguientes expresiones (Brereton, 2000):

$$x = T \cdot P + E \quad (25)$$

$$c = T \cdot q + f \quad (26)$$

Estas matrices se han representado en la figura 15. El producto de las matrices TP y Tq se aproxima a los valores medidos con la configuración instrumental utilizada y a las concentraciones reales de cada muestra, respectivamente E y f son las matrices de error. La matriz de puntuaciones T es común a ambos bloques, X y c .

Como se observa en la figura 12 para cada componente “a” existe un vector de puntuaciones o “scores”, un vector de pesos s o “loadings” y una matriz de residuales.

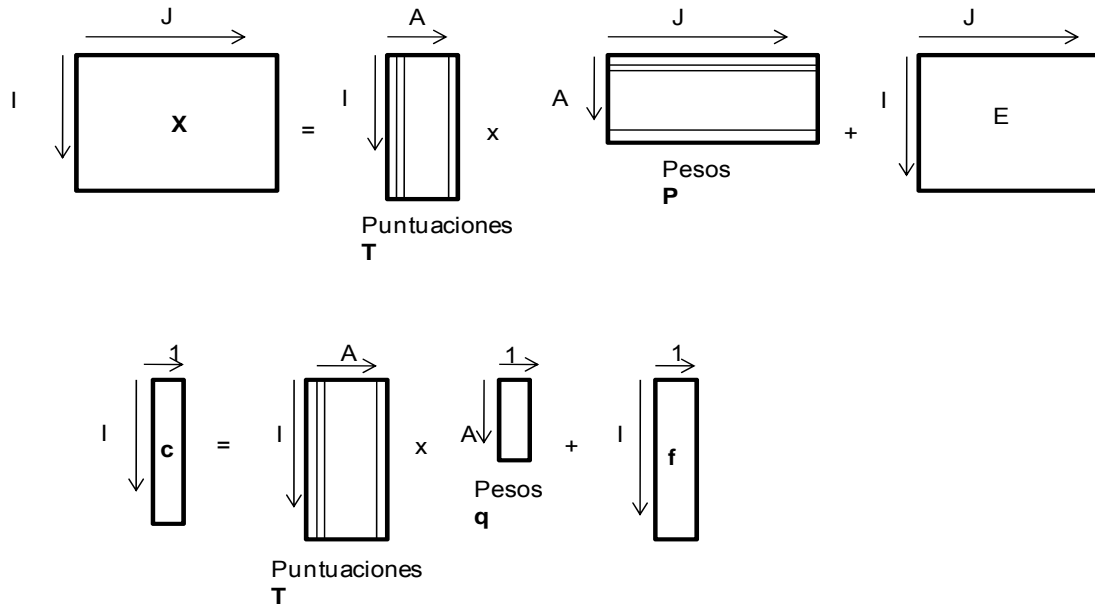


Figura 15. Matrices de PLS

Cada compuesto considerado genera una matriz de puntuaciones distintas, al contrario que en otras técnicas como PCR donde hay una única matriz de puntuaciones para todos los compuestos de la mezcla.

4.9.2.2. PLS2

La calibración mediante PLS2 permite el uso de una matriz de concentraciones C , en lugar de los vectores concentración c de cada compuesto individual de la mezcla. La ecuación (26) se transforma ligeramente y q se convierte en la matriz Q . el numero de columnas en C y en Q es igual al número de compuestos estudiados. En PLS1 se requiere que cada compuesto sea modelado independientemente mientras que en PLS2 todos se modelan conjuntamente. Las concentraciones estimadas mediante ambas técnicas pueden no ser iguales.

PLS2 puede ser útil (Brereton, 2003), en los casos donde existe correlación entre la concentración de los compuestos estudiados. En lugar de obtener un modelo independiente para cada compuesto de una mezcla, PLS2 permite analizar todos los datos de una vez. Sin embargo en muchas situaciones las predicciones realizadas

mediante esta técnica son peores que las estimaciones correspondientes mediante PLS1. Una posible estrategia puede ser realizada por PLS2 sobre el conjunto de datos como primera aproximación y después realizar PLS1 individualmente sobre cada compuesto conocido.

4.9.3. Validación de modelos

Uno de los aspectos más importantes de los métodos de calibración es la validación de los modelos. Esta facilita la información fundamentalmente sobre el número de componentes significativos necesarios para caracterizar un conjunto de datos y sobre la capacidad predictiva de los modelos en futuras muestras. Es posible obtener un buen ajuste entre los valores reales y los calculados mediante PLS aumentando el número de componentes considerados. Éste puede llegar, incluso, a ser completo si se utilizan todos los componentes PLS. Los principales métodos de validación se describen a continuación.

➤ Auto predicción

Es la aproximación más sencilla y se basa en calcular la raíz cuadrada de a media del error de calibración al cuadrado para cada compuesto, se calcula mediante la expresión:

$$E_{cal} = \sqrt{\frac{\sum_{i=1}^I (c_i - \hat{c}_i)^2}{I}} \quad (27)$$

Donde I es el número de muestras y la expresión entre paréntesis representa la diferencia entre la concentración puesta y predicha por el modelo. Generalmente este modelo de validación no se utiliza, ya que los datos pueden llegar a tener un ajuste perfecto aumentando simplemente el número componentes del modelo.

➤ Validación cruzada (cross validation)

La capacidad predictiva del modelo se comprueba separando un pequeño número de muestras del conjunto de calibración y analizando las predicciones obtenidas para este conjunto para el modelo reducido.

Aunque existen diversos criterios, es habitual retirar solo una muestra del modelo y predecir su concentración. Este se repite I veces de modo que al final todas las muestras se han predicho una vez. El error cometido cuando se utiliza la validación cruzada (CV) se calcula mediante la expresión:

$$E_{cv} = \sqrt{\frac{\sum_{i=1}^I (C_i - {}^{cv}\hat{C})^2}{I}} \quad (28)$$

donde ${}^{cv}\hat{C}_i$ es la concentración predicha de la muestra i del grupo de validación cruzada.

➤ Validación externa (test set)

Uno de los puntos débiles de la validación cruzada es que depende de cómo se ha realizado el diseño experimental para formar el modelo de calibración. Si dos compuestos están correlacionados, de forma que cuando la concentración de uno aumenta la del otro también lo hace y al revés, la auto predicción e incluso la validación cruzada puede sugerir que el modelo de la calibración es adecuado. Sin embargo si una futura muestra de predicción con un contenido elevado de los compuestos presenta un contenido bajo del otro es muy probable que los resultados de predicción se alejen de los valores reales. La validación cruzada es muy útil para informar sobre el ruido instrumental o errores de dilución y pesada por ejemplo, pero no es de gran ayuda si existe correlación en las concentraciones del grupo de patrones de calibración.

En lugar de la validación interna es posible comprobar la calidad predictiva del modelo con un conjunto independiente de muestra que no se han utilizado en ninguna etapa durante la construcción del modelo. El error cometido cuando se utiliza un grupo independiente o “test set” de I muestras es:

$$E_{test} = \sqrt{\frac{\sum_{i=1}^I (C_i - \hat{C}^{test})^2}{I}} \quad (29)$$

donde \hat{C}_i^{test} la concentración predicha de la muestra i del grupo de validación externa.

A partir de los valores absolutos de las ecuaciones (27-28) es habitual expresar el resultado en términos relativos, nuevamente, mediante la expresión (15).

Finalmente el sesgo (bias) del modelo y del grupo de validación se cuantifica mediante la expresión:

$$S = \frac{\sum_{i=1}^I (C_i - \hat{C})}{I} \quad (30)$$

Donde I es el número de muestras del modelo o del grupo de validación.

4.9. Aplicación de los métodos quimiométricos

Los métodos quimiométricos son empleados con mayor frecuencia, ya que son rápidos, confiables, económicos, efectivos y amigables con el medio ambiente (Irudarayaj *et al.*, 2001). En la industria de alimentos se utiliza con éxito en la caracterización de distintos productos (Downey, 1998) como aceite de oliva (Montoliu, 2001; García *et al.*, 2005) y vinos (Coimbra *et al.*, 2002; Edelman *et al.*, 2003). También en determinación de variedades de frutas como el caso de cultivares de fresa (Macías-Rodríguez *et al.*, 2004) así como en la determinación de adulteraciones y la correcta tipificación

geográfica de mieles (Gallardo-Velázquez, 2009; Reyes, 2009) y el jugo de manzana (Sivakesava *et al.*, 2001). La quimiometría es usada también en la tipificación de muestras de harinas de pescado, carne y soja (Cozzolino *et al.*, 2009).

Uno de los primeros campos de aplicación de la quimiometría fue el estudio de datos medioambientales. La concientización a nivel mundial sobre la importancia del estudio de la contaminación del medio ambiente, propició la realización de numerosos estudios que han generado un enorme volumen de datos, esto ha hecho que se recurra a disciplinas que, como la quimiometría, permiten su análisis y a la obtención e interpretación de la información sobre los procesos ambientales (Morer, 2003).

Hasta la fecha, pocos estudios han informado sobre el uso de dispositivos de análisis espectrales NIR para caracterizar suelos contaminados con petróleo. Malley *et al.*, (1999) informó de las relaciones de la regresión lineal entre las concentraciones de hidrocarburos totales de petróleo (TPH) predichas por NIR y los datos de referencia. Otros trabajos hacen uso de la tecnología para la estimación rápida de diesel en suelos contaminados (Malley *et al.*, 1999); en la determinación cuantitativa de las concentraciones totales de Zn, Pb y otras propiedades químicas y microbianas en suelos forestales contaminados con metales pesados (Chodak *et al.*, 2007); para predecir la cuantificación de hidrocarburos totales de petróleo contenidos en suelos contaminados (Chakraborty *et al.*, 2010); para la detección de petróleo y contaminación por combustibles como la gasolina, diesel, aceite de motor y mezclas sintéticas de hidrocarburos (Zwanziger y Forster, 1998). Así como para determinar parámetros cinéticos y termodinámicos de la adsorción de asfaltenos en cuarzo, dolomita, arena y mica (Syunyaev *et al.*, 2009).

Recientes investigaciones han demostrado las ventajas de la espectroscopía para predecir en forma rápida y no destructiva las propiedades físicas, químicas y biológicas del suelo (Janik *et al.*, 1998; Jarquín, 2006). Aunque la técnica NIRS es una herramienta muy promisoriosa en otras disciplinas, no ha recibido mucha atención en el campo de la ciencia del suelo (Ben-Dor y Banin, 1995).

También el desarrollo de la estadística ha sido fundamental en el progreso de esta tecnología como método de análisis. Con la aplicación del análisis estadístico de regresión a la reflectancia difusa, la tecnología NIRS ha evolucionado enormemente, pasando a ser hoy en día, un potente sensor para el análisis cualitativo y cuantitativo en la industria agroalimentaria, farmacéutica, química y en determinadas aplicaciones en medicina, medioambiente, entre otras (Davies y Williams, 1996).

La contribución del presente trabajo está enfocada a la aplicación de los métodos quimiométricos acoplado a la tecnología NIR en el área de contaminación ambiental, con la finalidad de contribuir con información que sirva de base para el uso futuro de esta técnica en esta área. El creciente uso de la quimiometría se ha demostrado en muchas investigaciones, pero como lo han citado algunos autores ha recibido poca atención en el campo de estudio del suelo y medioambiental. Son indiscutibles las bondades de la técnica quimiométrica sobre todo durante la manipulación y análisis de las muestras, las cuales no sufren alteraciones, quedando disponibles para otros estudios. Este trabajo plantea la posibilidad de uso de la técnica NIR en la predicción de forma rápida y no destructiva de contaminantes presente en el suelo, así como el seguimiento de su proceso de degradación.

5. MATERIALES Y METODOS

El desarrollo experimental se realizó en dos etapas. En la primera se llevó a cabo un muestreo de suelo en campo no afectado por contaminación de petróleo. En la segunda etapa el suelo llevado al laboratorio fue contaminado artificialmente con petróleo crudo en una concentración de 50000 mg kg^{-1} de HTP, a partir del cual se llevaron a cabo dos actividades experimentales. La primera actividad consistió en cuantificar los cambios en la concentración de petróleo crudo por efecto de degradación microbiana nativa en condiciones controladas de humedad y temperatura, mediante espectroscopía de infrarrojo (NIRS). Como segunda actividad experimental y de manera alterna se realizaron extracciones por reflujo en Soxhlet de acuerdo al método 3540C de la USEPA (1996), técnica analítica utilizada como método de referencia.

5.1. Etapa de Campo

5.1.1. Localización del área de muestreo

El muestreo del suelo Gleysol se realizó en el área de la Cuenca Baja del río Tonalá y lagunas del Carmen y Machona que cuenta con una superficie de 109,000 ha, localizada en el sureste de México, al oeste del estado de Tabasco en los municipios de Cárdenas y Huimanguillo. El sitio se sitúa en la provincia fisiográfica de llanuras y pantanos tabasqueños, con clima cálido húmedo (Amf), temperatura media anual de 26 °C y precipitación media anual de 2200 mm. La vegetación típica es selva baja perennifolia, manglares, vegetación hidrófila y palmar (Zavala, 2009) (Figura 16).

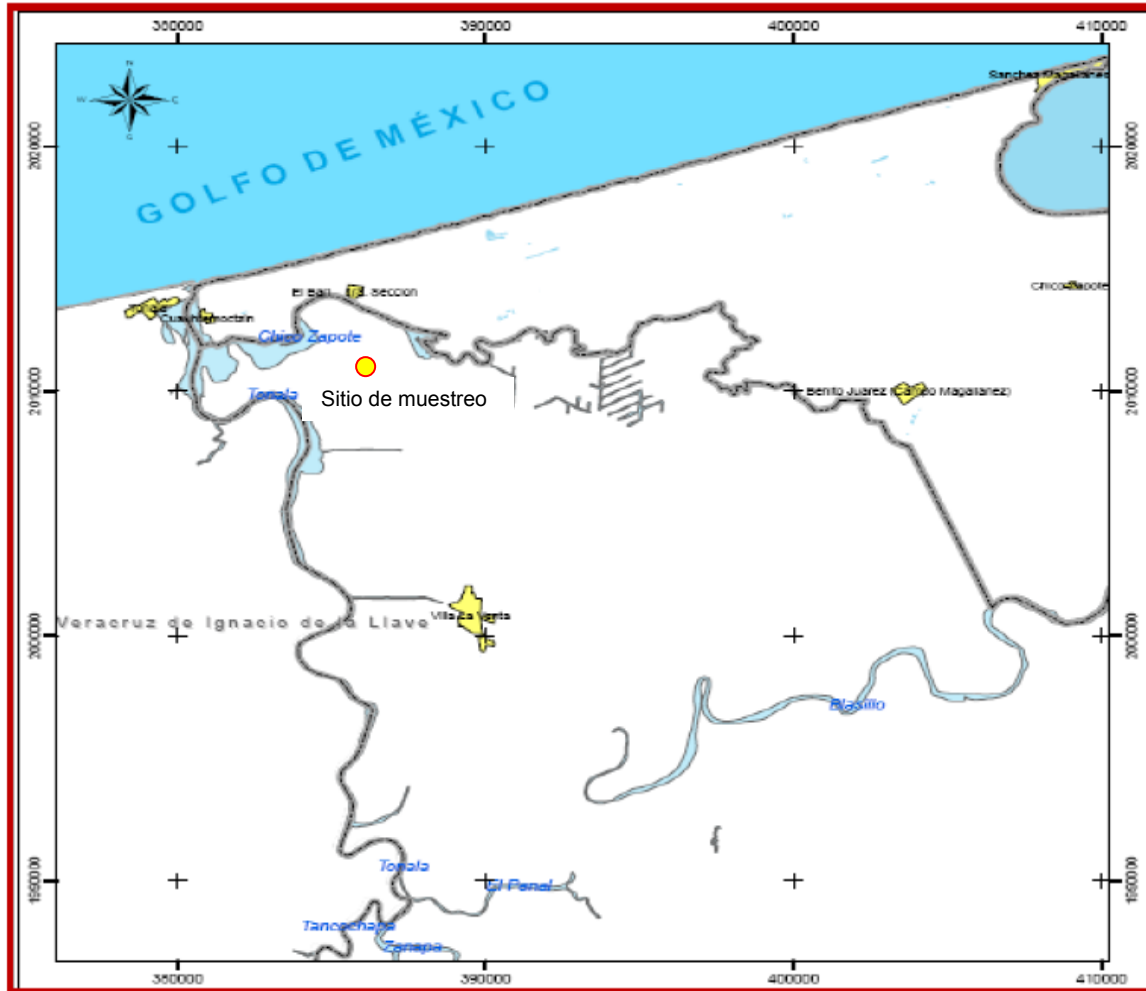


Figura 16. Ubicación Geográfica del sitio de muestreo del suelo Gleysol, en el Ejido Aquiles Serdán, La Venta, Tabasco.

5.1.2. Muestreo del suelo Gleysol

El muestreo de suelo en campo fue dirigido, como lo define la Norma Oficial Mexicana (NOM-138-SEMARNAT/SS-2003). Este muestreo se realizó en el Ejido Aquiles Serdán, de La Venta, Tabasco, entre las coordenadas geográficas $18^{\circ} 10' 43.38''$ latitud norte y $94^{\circ} 04' 12.08''$ longitud oeste. Se tomaron aproximadamente 60 kg de suelo a una profundidad de 30 cm con una pala recta limpia (Figura 17).



Figura 17. Sitio de muestreo del suelo Gleysol (0-30 cm), Ejido Aquiles Serdán, de La Venta, Tabasco.

En este sitio la vegetación característica esta representada por las familias Fabaceae, Poaceae y Cyperaceae. En el sitio se localizan árboles de apompo (*Pachira aquatica*), anonillo (*Annona glabra*), estribo (*Dalbergia brownei*) e icaco (*Chrysobalanus icaco*); en el estrato herbáceo hay helechos como guanillo (*Blechnum serrulatum*) y *Thelypteris interrupta*, pasto pelillo (*Leersia hexandra*), *Begonia fischeri*, *Scleria microcarpa* y *Scleria macrophylla*. El matorral inundable esta representado por el tasiste (*Acoelorrhaphe wrightii*) (García-López *et al.*, 2006) (Figura 18).



Figura 18. Vegetación representativa del sitio de muestreo, Ejido Aquiles Serdán, La Venta, Tabasco.

5.1.3. Características del suelo Gleysol (GL)

Los Gleysoles son suelos de humedales, que a menos que sean drenados, están saturados con agua freática por periodos suficientemente largos para desarrollar un patrón característico de color gléyico (procesos de reducción o una reducción asociada a la segregación de hierro, dichos procesos se observan por la presencia de colores verdosos o azulosos) dentro de los 50 cm de profundidad (FAO, 1990; FAO, 2007; Palma-López *et al.*, 2007). Zavala *et al.* (2003), consideran que este tipo de suelo es uno de menos estudiados en Tabasco, debido al poco interés agronómico que presenta y sobre todo a las dificultades de campo para su estudio, localmente se le conoce como popalerías o bajiales.

5.2. Etapa de Laboratorio

5.2.1. Preparación de las muestras

El suelo fue secado a la sombra durante una semana a temperatura ambiente, y luego

molido con rodillo de madera y tamizado en malla de 2 mm de diámetro. Posteriormente el suelo fue depositado en un tambo de PVC con capacidad de 20 kg y sometido a homogenización rotatoria durante un periodo de 20 horas.

5.2.2. Características físicas y químicas del suelo

A la muestra de suelo se le realizó una caracterización de los principales parámetros físicos y químicos. Estas determinaciones fueron humedad, pH, materia orgánica (MO), conductividad eléctrica (CE), nitrógeno total (NT), capacidad de intercambio catiónico (CIC), bases intercambiables y textura. Los análisis fueron realizados con los métodos establecidos en la NOM-021-RECNAT-2000.

5.2.2.1. Determinación de humedad

La humedad de un suelo se define como la cantidad de agua que existe en un suelo, referida en porcentaje. La humedad se determina por métodos gravimétricos comúnmente, es decir, una muestra húmeda se pesa, se seca a la estufa a 105 °C y se vuelve a pesar, obteniéndose por diferencia la masa de agua (Etchevers *et al.*, 1992). Esta masa de agua es referencia de la masa de suelo seco de la muestra. La determinación del contenido de humedad del suelo por gravimetría, se realizó a través del método AS-05 de la NOM-021-RECNAT-2000.

5.2.2.2. Determinación de pH

La determinación del pH en suelo medido en agua se realizó en una suspensión de suelo:agua en proporción 1:2 (p/v). Para ello se resuspendieron 10 g de suelo en 20 ml de agua, se agitó la mezcla durante 30 minutos y se midió el pH en un potenciómetro Orion 710A+ Thermo Electron Corporation, previamente calibrado con soluciones amortiguadoras de pH 4 y 7. La determinación del potencial de hidrógeno se llevó a cabo por el método AS-02 de la NOM-021-RECNAT-2000.

5.2.2.3. Determinación de materia orgánica

Esta determinación se basa en cuantificar el carbono orgánico del suelo que se oxida con dicromato de potasio en presencia de ácido sulfúrico. El exceso de oxidante se valora con sulfato ferroso y la cantidad de carbono orgánico oxidado se calcula a partir de la cantidad de dicromato reducido. La determinación de la materia orgánica se llevó a cabo de acuerdo al método AS-07 (Walkley y Black) de la NOM-021-RECNAT-2000.

5.2.2.4. Determinación de textura

El método utilizado para obtener la textura del suelo fue el de Bouyucos que se basa en las diferentes velocidades de sedimentación del particulado según la ley de Stokes, utilizando la técnica de la sedimentación según Gee y Bauder (1986). El método permitió cuantificar las fracciones de arena, limo y arcilla en forma aproximada las cuales son suficientes para determinar posteriormente la textura. El método utilizado fue el AS-09 de la NOM-021-RECNAT-2000.

5.2.2.5. Determinación de la conductividad eléctrica

La conductividad eléctrica es la capacidad de una solución acuosa para transportar una corriente eléctrica. La determinación de la conductividad eléctrica es por lo tanto una forma indirecta de medir la salinidad del agua o extractos de suelo. Este método se basa en la teoría de la disociación electrolítica (Fernández *et al.*, 2006).

Para calcular la conductividad eléctrica (CE) del suelo (parámetro relacionado con la presencia de iones en el suelo) se partió de una resuspensión suelo:agua en proporción 1:5 (p/v) después de 24 horas se midió la conductividad ($\mu\text{S cm}^{-1}$) de la resuspensión con un conductímetro YSI 3200. Para esta determinación se uso el método AS-18 de la NOM-021-RECNAT-2000.

5.2.2.6. Determinación de nitrógeno total

Dada la dinámica del nitrógeno dentro del suelo, su determinación solo informa la cantidad disponible en el momento de realizar el análisis. La determinación de nitrógeno involucra la digestión de la muestra para convertir el Nitrógeno a NH_4^+ y la determinación de NH_4^+ en el digestado. La digestión de la muestra es desarrollada por calentamiento de la muestra con H_2SO_4 concentrado, K_2SO_4 y CuSO_4 que promueven la oxidación de la materia orgánica y la conversión del nitrógeno orgánico a amonio por incremento de la temperatura de digestión y el empleo de catalizadores. El amonio en el digestado es determinado por titulación del amonio liberado por destilación del digestado con álcali. Para la determinación del nitrógeno total en el suelo contaminado se realizará de acuerdo al método AS-25 de la NOM-021-RECNAT-2000.

5.2.2.7. Determinación de la capacidad de intercambio catiónico y bases intercambiables

La CIC o capacidad de intercambio catiónico es la capacidad del suelo para retener e intercambiar diferentes elementos minerales. Esta capacidad aumenta notablemente con la presencia de materia orgánica, y podría decirse que es la base de lo que se llama fertilidad del suelo.

Para la determinación de la capacidad de intercambio catiónico (CIC) y bases intercambiables (Na^{2+} , Mg^{2+} , Na^+ y K^+) del suelo, se empleó el acetato de amonio como solución saturante. El método para la determinación consiste en la saturación de la superficie de intercambio con un catión índice, el ión amonio; lavado del exceso de saturante con alcohol; desplazamiento del catión índice con potasio y determinación del amonio mediante destilación. El amonio se emplea como catión índice debido a su fácil determinación, poca presencia en los suelos y porque no precipita al entrar en contacto con el suelo. La determinación de la capacidad de intercambio catiónico y bases intercambiables se hizo de acuerdo al método AS-12 de la NOM-021-RECNAT-2000.

5.3. Análisis de HTP mediante el método Soxhlet y la tecnología NIR

5.3.1. Determinación de los HTP en las muestras de suelo contaminado, por el método convencional Soxhlet

A partir de la incubación de los tratamientos con petróleo crudo y sin petróleo crudo, se evaluó el contenido del hidrocarburo por el método convencional Soxhlet. La extracción y la cuantificación de los HTP se realizó por reflujo utilizando el método 3540C (USEPA, 1996). La extracción se efectuó con diclorometano (Baker) en un equipo Soxhlet. El método consistió en colocar diez gramos de suelo seco en un cartucho de extracción de celulosa, al cual se le adicionaron 10 g de Na_2SO_4 anhidro en una relación suelo:sulfato 1:1, las muestras fueron colocadas en la columna extractora del equipo Soxhlet, se le adicionó 130 mL de CH_2Cl_2 y se procedió a la extracción del petróleo crudo de la matriz del suelo por un periodo de ocho horas (Figura 19). Las determinaciones de HTP se realizaron en las mismas fechas que las lecturas con el equipo NIRS; en los días 0, 1, 2, 4, 8, 16, 32.



Figura 19. Tren de extracción de HTP en equipo Soxhlet.

5.3.2. Determinación de HTP en las muestras de suelo contaminado por espectroscopía de infrarrojos NIR

Para realizar el análisis con el NIRS, las muestras fueron depositadas en las celdas de transporte de muestra para registrar los espectros. Los datos obtenidos de las lecturas con NIRS y del método convencional fueron analizados por quimiometría para generar los modelos matemáticos de predicción para el proceso de degradación del petróleo crudo en el tratamiento contaminado.

El equipo utilizado para el análisis fue un espectrómetro de Reflectancia en el Infrarrojo Cercano modelo Versatile Agri Analyzer™ 6500 System II Vertical Type 461008, Serial No. 9503 100-240 Vac max 200 W, 50-60 Hz de la empresa FOSS NIRSystems (Figura 20), el intervalo que cubre el escáner es de 400 a 2500 nm.



Figura 20. Instrumento NIR Versatile Agri Analyzer™ 6500.

Para realizar la lectura el equipo cuenta con una lámpara de 4 Volts, ajustada a una placa prealineada, que manda el haz de luz blanca la cual contiene todas las frecuencias al *Grating*, una superficie holográfica de 570 líneas por mm y una distancia

de 1.75 μm entre líneas, el cual desvía la luz en ángulos diferentes dependiendo de la longitud de onda, ésta luz reflejada es monocromada y el ángulo de inclinación del *Grating* es controlado por la tarjeta madre, al filtro *Order sorter*, que esta sincronizado con el movimiento del *Grating* y es movido por medio de un motor de pasos, permite el paso de la luz por longitud de onda para transmitirla a la muestra, contenida en una copa o celda opaca de cuarzo, y esta a los detectores de silicio (400 a 1100 nm) y PbS (1100 a 2500 nm), que registran la cantidad de luz reflejada y/o transmitida, para transformarla en una señal que se envía mediante el programa ISIScan (versión 2.81 Copyright, 2006) a una computadora (Acer AMD Sempron™ Processor 3200+ 1.81 GHz, 448 MB de RAM). Los datos de reflectancia se guardaron como el logaritmo ($1/R$) (donde R = reflectancia) a 2 nm de intervalo (1.050 datos por muestra).

5.4. Análisis Quimiométrico de los Resultados

Los datos obtenidos como resultado del análisis químico de extracción Soxhlet y de las lecturas en el espectrofotómetro NIR se analizaron quimiométricamente, utilizando los programa V-PARVUS versión 2008 y Pirouette Chemometrics Software versión 3.11 para Windows (Infometrix, 2003). Se aplicaron las técnicas de reconocimiento no supervisadas para determinar relaciones entre muestras sin utilizar ninguna información previa con respecto a la pertenencia a determinadas clases. Permitiendo conocer la existencia de estructuras latentes en los datos y obtener información útil contenida en los mismos. El análisis aplicado fue el de componentes principales (PCA) que permite encontrar la estructura “latente” en la matriz de datos cuya lectura directa lo impide; reduciendo la dimensión del conjunto de datos, construyendo nuevas variables, ortogonales entre sí, que son combinaciones lineales de las variables originales.

Así mismo, se aplicaron técnicas de reconocimiento de patrones supervisados, para clasificación como el análisis discriminante lineal (LDA), que nos permite la creación de una regla de clasificación entre clases. LDA es una técnica de aprendizaje supervisado para clasificar datos. La idea central de LDA es obtener una proyección de los datos en un espacio de menor (o incluso igual) dimensión que los datos entrantes, con el fin de

que la separación de las clases sea la mayor posible. Es una técnica supervisada ya que para poder buscar esa proyección se debe entrenar el sistema con patrones etiquetados (García, 2009).

Se aplicó también el modelado suave independiente de analogía de clases (SIMCA), que es una técnica modelativa, paramétrica y probabilística que consiste en desarrollar un modelo matemático, en este caso es un modelo de componentes principales, que defina un dominio en el espacio para cada muestra del conjunto de entrenamiento (Infometrix, 2003). Disponiendo de información previa como el número de componentes principales utilizados para cada clase o manteniendo el número de ellos que contengan la máxima información discriminante. Finalmente a través de una regresión por mínimos cuadrados parciales (*PLS*, por sus siglas en inglés), se obtuvo el modelo para predecir el fenómeno de degradación del petróleo crudo en el suelo Gleysol.

5.5. Evaluación de la Población de Hongos Filamentosos Nativos

Para llevar a cabo el experimento de degradación del petróleo crudo, se trabajó con la población de hongos filamentosos nativos del suelo. Esto con la finalidad de conocer la pérdida de HTP por biodegradación por los hongos nativos.

El recuento del tamaño de la población de los hongos nativos se realizó para los días 4, 8, 16, 32, utilizando el método de cuenta viable en dilución seriada de acuerdo con Raina *et al.*, (2000). Para el conteo de hongos totales se utilizó el medio de cultivo Papa Dextrosa Agar (Bioxon).

5.6. Características del Petróleo Crudo

Para este trabajo se utilizó el petróleo crudo superligero “Olmeca”, con el cual se contaminó el suelo Gleysol a una concentración de 50000 mg kg⁻¹. Las características de este crudo se presentan en el cuadro 3.

Cuadro 3. Propiedades del petróleo crudo “Olmeca”.

Propiedades	Olmeca
Peso específico a 20/4 °C	0.8261
Grado API	39.16
Viscosidad SSU	
15.6 °C	43.3
21.1 °C	40.3
25.0 °C	39.0
Azufre total, % peso	0.81
Carbón Ramsbottom, % peso	1.62
Cenizas, % peso	0.006
Aceites, % peso	89.2
Parafinas, % peso	13.4
Aceite desparafinado, % peso	75.8
Resinas, % peso	10.8

5.7. Diseño experimental

Se estableció un bioensayo *In Vitro*, basado en un diseño experimental completamente al azar (DCA) (Camacho *et al.*, 1992; Reyes, 2003). El tratamiento uno fue contaminado con 50000 mg kg⁻¹ de HTP, al tratamiento dos no se le aplicó petróleo crudo, ya que sirvió como suelo testigo para realizar una comparación del tamaño de las poblaciones de hongos filamentosos con respecto al tratamiento uno, durante el proceso de degradación (Cuadro 4). En cuanto a los microorganismos para determinar la degradación del petróleo crudo, se trabajó con la población de hongos filamentosos nativos del suelo Gleysol.

Cuadro 4. Descripción de tratamiento y unidades experimentales

Tratamiento	Descripción	Unidades Experimentales						
1	Gleysol + 50000 mg kg ⁻¹ de HTP*	4	3	8	1	6	11	
2	Gleysol + 0 mg kg ⁻¹ de HTP*	9	12	2	5	10	7	

* Ambos tratamiento fueron evaluados únicamente con la flora microbiana nativa (hongos filamentosos)

5.8. Establecimiento del Experimento

Los contenedores utilizados fueron macetas de aluminio de un kilogramo cubiertas con

plástico poroso (marca *Kleen Pack*), el cual permite el intercambio de aire, pero no de partículas. En cada maceta se colocaron 500 g de suelo Gleysol, el cual se mezcló con 50000 mg kg⁻¹ del petróleo crudo Olmeca. La cantidad de agua que se adicionó a los tratamientos fue hasta capacidad de campo la cual se mantuvo durante todo el periodo del bioensayo.

Los contenedores se colocaron en una estufa de incubación a 35 °C durante 30 días. Se aplicó riego a capacidad de campo cada dos días para compensar la pérdida de humedad por evapotranspiración. Durante este tiempo se realizaron evaluaciones periódicas para evaluar la biodegradación con las técnicas propuestas.

5.9. Análisis Estadístico de la Población de Hongos Filamentosos

Se realizó un análisis de varianza del diseño experimental (DCA) para las poblaciones de hongos en ambos tratamientos, así como para determinar diferencias con respecto al porcentaje de degradación de petróleo crudo por tiempo evaluado. Igualmente se hizo una prueba de comparación de medias (Tukey), para comprobar las diferencias entre tratamientos y tiempo evaluado. Los datos fueron analizados con el procedimiento PROC GLM del programa Statistical Analysis System (SAS, 2002).

6. RESULTADOS Y DISCUSIÓN

La evaluación del porcentaje de pérdida de petróleo crudo por espectroscopía de infrarrojo, después de un proceso de volatilización y/o degradación por microorganismos nativos del Gleysol durante treinta y dos días de experimentación, fue con la finalidad de obtener información suficiente, interpretarla y colocarla en un contexto apropiado para convertirlos a través de la quimiometría en información útil.

6.1. Propiedades Físicas y Químicas del Suelo

La fertilidad y el funcionamiento de los suelos dependen en gran medida de las propiedades bioquímicas y microbiológicas, ya que son muy importantes para definir las principales funciones edáficas como la productividad o la degradación. Por lo tanto, la actividad biológica y bioquímica del suelo reviste gran importancia en el mantenimiento de la fertilidad de los hábitats terrestres.

Los análisis físicos y químicos se realizaron para caracterizar los parámetros más importantes de la muestra de suelo, estos parámetros fueron pH, materia orgánica (MO), textura, conductividad eléctrica (CE), nitrógeno total (NT), capacidad de intercambio catiónico (CIC) y bases intercambiables, de acuerdo a la observancia de la NOM-021-RECNAT-2000. Las propiedades físicas y químicas se presentan en el Cuadro 5.

6.1.1. Determinación de la humedad

La humedad del suelo Gleysol medida en base a la norma fue de 13.8 %, encontrándose por debajo del promedio que según Ortiz (1999) es de 28 a 32 %. De acuerdo a este autor uno de los factores importantes en el suelo es la humedad, ya que a bajas o altas cantidades de agua el crecimiento microbiano se inhibe, mientras que a porcentajes medio el suelo presenta oxígeno, permitiendo que las poblaciones de microorganismos afloren.

Cuadro 5. Propiedades físicas y químicas de la muestra de suelo Gleysol no contaminado (NOM-021-RECNAT-2000).

Parámetros		Método analítico	Resultados
Humedad (%)		Gravimetría (AS-05)	13.8
pH (H ₂ O)		Potenciómetro (AS-02)	7.0
MO (%)		Walkley y Black (AS-07)	6.37
CE (μS cm ⁻¹) (H ₂ O)		Conductímetro (AS-18)	692.2
NT (%)		Digestado (AS-25)	0.18
CIC			33.5
K ⁺			0.65
Ca ²⁺	Cmol (+) kg ⁻¹	Acetato de amonio (AS-12)	16.47
Na ⁺			2.35
Mg ²⁺			9.87
Textura			
	Arena (%)		29
	Limo (%)	Bouyoucos (AS-09)	30
	Arcilla (%)		43
			Arcillosa

El agua es esencial para todos los seres vivos porque participa en varias reacciones metabólicas celulares, actúa como un solvente y portador de nutrimentos desde el suelo hasta las plantas y dentro de ellas. La solución del suelo contiene un determinado número de sales inorgánicas, que son importantes para la flora microbiana (Alexander, 1990).

La actividad microbiana se ve beneficiada en un potencial de agua -0.1 atm o capacidad de campo. En potenciales altos, la ventilación constituye un problema, ya que la difusión del O₂ resulta obstaculizada (Coyne, 2000). Por su parte, Brzezińska *et al.*, (1998), señala que el incremento en el agua del suelo reduce el estado de aireación por reducción del espacio de los poros llenos de aire disponibles para la difusión de gases, por lo que la elevación de la humedad del suelo en determinado nivel puede modificar considerablemente el equilibrio microbiano (Fenche *et al.*, 2000). En el extremo opuesto

Jenkinson (1992), considera que cuando el potencial hídrico alcanza valores muy negativos, la actividad microbiana cesa. Es indiscutible por tanto que la disponibilidad de agua es un factor importante ya que regula las actividades físicas, químicas y biológicas en el suelo.

6.1.2. Determinación del pH

El pH es una propiedad química del suelo que tiene un efecto importante en el desarrollo de los seres vivos, tanto microorganismos como plantas. La lectura de pH se refiere a la concentración de iones hidrógeno activos (H^+) que se da en la interfase líquida del suelo, por la interacción de los componentes sólidos y líquidos. La concentración de iones hidrógeno es fundamental en los procesos físicos, químicos y biológicos del suelo (Fernández *et al.*, 2006).

El resultado de pH para este suelo fue de 7.1 que de acuerdo con la norma mexicana indica un suelo neutro (Cuadro 5), ya que se ubica en el rango de 6.6 a 7.3. El pH del suelo es importante para el desarrollo de los microorganismos degradadores de la materia orgánica, siendo los más adecuados los comprendidos entre 6 y 8. Para las bacterias mientras mayor sea la concentración de iones de hidrógeno, menor es el tamaño de sus poblaciones, por lo que el incremento del pH como el encalado, aumenta apreciablemente la abundancia de estos microorganismos. Para el caso de los hongos, pueden dominar la comunidad microbiana cuando el pH es bajo, es decir, son acidotolerantes. Esto no es consecuencia de que los hongos encuentren su óptimo en condiciones ácidas, si no que es el resultado de la ausencia de la competencia de otros microorganismos por las reservas alimenticias. De este modo los hongos son los responsables de una parte considerable de las transformaciones bioquímicas en hábitats ácidos (Alexander, 1990; Atlas y Bartha, 2002). En suelos ácidos los hongos son más competitivos que las bacterias, las cuales prefieren ambientes neutros. Por lo tanto, en suelos con bajo pH, los hongos se vuelven más importantes en la degradación de compuestos orgánicos, incluyendo los hidrocarburos (Raina *et al.*, 2000).

El pH puede afectar de dos maneras a las poblaciones y actividad de los microorganismos del suelo. Indirectamente, cuando el pH es bajo, los ácidos orgánicos son tóxicos para el desarrollo microbiano. Esta condición se agrava si existe carencia de oxígeno, pues la materia orgánica se acumula y la acidez del suelo aumenta. El efecto directo del pH se presenta por la afinidad de los organismos (bacterias y hongos) a diferentes condiciones; también por la afinidad de los procesos catalíticos, como la formación de nitratos, conocido como nitrificación, que se da en rangos de pH determinados (Tate, 2000). Acuña *et al.* (2006), menciona que a un pH del suelo medianamente ácido, una acidez intercambiable intermedia, altos contenidos de materia orgánica y alta disponibilidad de elementos esenciales favorece la actividad de los hongos. El pH del suelo también puede afectar la solubilidad del fósforo, un nutriente importante para los microorganismos, así como el transporte de los contaminantes (Eweis *et al.*, 1999).

De acuerdo a lo anterior y a la capacidad de la mayoría de los microorganismos del suelo, el pH del Gleysol favorece el adecuado crecimiento y actividad biológica de la población de hongos filamentosos presentes en él, aún cuando estuvieron expuestos al petróleo crudo, ya que este no tiene ningún efecto sobre el pH (Martínez y López, 2001).

6.1.3. Determinación de la materia orgánica

La materia orgánica (MO) del suelo evaluado a través del contenido de carbono orgánico por el método Walkley y Black, que se basa en la oxidación del carbono orgánico del suelo, presentó un 6.39 % de materia orgánica, porcentaje alto de acuerdo con la norma mexicana (Cuadro 5). Jarquín (2006), menciona que a mayor contenido de materia orgánica los suelos tienden a ser más oscuros, retener mayor humedad, la consistencia se vuelve más friable y por lo general estos suelos son utilizados con fines agrícolas en el trópico.

La materia orgánica del suelo es la fracción orgánica que incluye residuos vegetales y

animales en diferentes estados de descomposición, incluye tanto a los materiales poco alterados como a aquellos que han experimentado cambios de descomposición, transformación y resíntesis dentro del suelo. Se pueden incluir también compuestos orgánicos tóxicos, provenientes de las actividades industriales del hombre, como la contaminación de suelos por hidrocarburos del petróleo, que también constituye parte de la materia orgánica del suelo, ya que esta facilita los mecanismos de absorción de las sustancias peligrosas (Etchevers, 1988; Julca *et al.*, 2006). La materia orgánica es uno de los componentes del suelo, ya que es una importante fuente de energía y nutrimentos para microorganismos como bacterias, hongos, protozoarios y nemátodos (Doran y Parkin, 1994).

También la materia orgánica tiene otra función importante en el suelo y esta relacionada con su carga negativa, pues tiene la propiedad de adsorber y retener a los microorganismos de manera significativa (Goyal y Gerba, 1979).

Mustin (1987), indica que la materia orgánica representa del 95 al 99% del total del peso seco de los seres vivos, pero su presencia en los suelos suele ser escasa y son contadas las excepciones en las que supera el 2% (Navarro *et al.*, 1995). Para Gros y Domínguez (1992), el nivel deseable de materia orgánica en los suelos arcillosos medios es del 2%. Con base en lo anterior y las consideraciones de la NOM-021-RECNAT-2000 en vigor, se tiene que el suelo Gleysol presentó un porcentaje alto de materia orgánica, el cual aumentó en forma proporcional a la concentración de los hidrocarburos de petróleo agregado durante la prueba de degradación en este trabajo. Martínez y López (2001), mencionan que la materia orgánica puede incrementar como consecuencia de la suma de la materia orgánica (biogénica) y la aportada por los hidrocarburos (petrogénicos).

El contenido de materia orgánica de un suelo es un buen indicador de su fertilidad, principalmente de su capacidad potencial para proporcionar nutrientes como nitrógeno, fósforo, azufre, etc. (Rodríguez y Rodríguez, 2002). La materia orgánica es el conjunto de residuos vegetales y animales, más o menos descompuestos por la acción de los

microorganismos del suelo, que se encuentra en estado de evolución en el suelo (Alexander, 1990). La determinación de la materia orgánica de los suelos puede referirse al total o solamente al humus estable, siendo este último el más significativo.

En este trabajo el incremento de la materia orgánica para el tratamiento contaminado con petróleo crudo, se dio por la adición de material de origen petrolífero. Esto puede representar un problema de toxicidad para la microflora nativa, sobre todo por la presencia de fracciones aromáticas. Aunque existen microorganismos como bacterias y hongos capaces de utilizar estos hidrocarburos como fuente de carbono y energía, otros carecen de esa capacidad para metabolizar compuestos petrogénicos. Estas excepciones por lo regular se presentan en suelo con flora microbiana que no está en presencia del contaminante para que desarrollen esa capacidad. Contrario a lo que mencionan algunos autores no siempre el aumento de materia orgánica por agregación intencional o no de material petrogénico es benéfico, principalmente cuando la concentración del contaminante se hace cada vez mayor, lo que puede representar un riesgo ecotóxico (Martínez y López, 2001).

6.1.4. Determinación de la textura del suelo

De acuerdo a los resultados, el suelo Gleysol utilizado en este trabajo tiene 29 % de arena, 30 % de limo y 43 % de arcilla (Cuadro 5), estos datos interpretados a través del triángulo de texturas, indican que el suelo Gleysol tiene una textura arcillosa.

De acuerdo con Razo (2003), la textura es una característica básica e influye en las poblaciones microbianas presentes en el suelo, así como en el tiempo de tránsito de contaminantes. Para que se lleve a cabo la actividad microbiana en el suelo es importante la formación de agregados, ya que las arcillas son los componentes básicos para que esto suceda, la mayoría de las arcillas tienen carga negativa, al igual que la superficie de las bacterias, y si hay alto porcentaje de arcillas, las partículas en el suelo se acomodan ajustadamente dejando poco espacio para que fluya el agua, lo que dificulta el manejo del suelo.

La formación de estos agregados es de interés microbiológico, ya que el material celular y excreciones de la flora microbiana son factores que afectan a la formación y estabilidad de los agregados. Así mismo, la estabilidad de estos agregados puede variar de acuerdo a ciertas condiciones como el manejo del suelo, actividad microbiana y otros factores (Alexander, 1990). En suelos arcillosos como el usado en este trabajo, los poros son generalmente pequeños, por lo que estos espacios afectan el movimiento y retención de agua haciéndolo más pesado, esta condición regula de forma importante la actividad óptima de la microflora del suelo. En este sentido los microorganismos se ven afectados en forma indirecta, ya que una textura pesada provee condiciones de reducción constantes lo cual afecta el crecimiento normal de la microflora del suelo, así como la disponibilidad de agua y aire para las raíces de las plantas (Cerisola *et al.*, 2005).

6.1.5. Determinación de la conductividad eléctrica

La medida de la conductividad eléctrica (CE) del suelo permite estimar la cantidad de sales que contiene. El análisis de la CE en suelos se hace para establecer si las sales solubles se encuentran en cantidades que permitan la germinación normal de las semillas, el crecimiento de las plantas o la absorción de agua por parte de las mismas (USDA, 1996). El aumento de la CE en el suelo puede provocar una disminución en la productividad y también es una limitante para el buen desarrollo de las actividades de los microorganismos en el suelo (Sabirova, 2000), pero a su vez puede originar en los organismos posibles adaptaciones a esas condiciones (Luna *et al.*, 2002).

La conductividad eléctrica en el Gleysol no contaminado fue de $692 \mu\text{S cm}^{-1}$ y de acuerdo con la NOM-021-RECNAT-2000, indica efectos despreciables de la salinidad sobre este suelo, es decir que la salinidad no representa un riesgo para las poblaciones microbianas nativas (Cuadro 5).

6.1.6. Determinación del nitrógeno total

El nitrógeno es un elemento indispensable para la vida, forma parte de las principales biomoléculas de todos los seres vivos. El nitrógeno puede llegar al suelo en los aportes de materia orgánica y a la fijación bacteriana a partir del aire. Dentro del suelo es aprovechado por las plantas, animales y microorganismos que lo incorporan a sus tejidos (Fernández *et al.*, 2006).

En los microorganismos la carencia de nitrógeno puede afectar su crecimiento, por lo que la población microbiana no tendrá un desarrollo óptimo. El nitrógeno es necesario para la descomposición de la materia orgánica por los microorganismos heterótrofos del suelo. Un valor óptimo de C:N para una descomposición rápida oscila entre 15:1 y 25:1. Sin embargo, con rangos mayores a 25:1 puede resultar que el nitrógeno quede inmovilizado por los microorganismos del suelo, ya que los mismos necesitarán de este nitrógeno para su nutrición (Tisdale *et al.* 1993), o en su caso pueden utilizar el amonio o nitratos presentes en el suelo. Este nitrógeno permite el rápido crecimiento de los microorganismos que proporcionan material con carbono al suelo, por lo tanto un aumento de nitrógeno trae consigo una mayor mineralización del C y N de la materia orgánica del suelo (Font, 2003).

Aun cuando su utilidad como indicador de disponibilidad de nitrógeno en el suelo es escasa, su determinación resulta necesaria en estudios de suelos. Su escasa utilidad es debido a que la mayor parte del nitrógeno se encuentra en forma orgánica con muy bajas tasas de mineralización. De acuerdo con los resultados del análisis el NT en el Gleysol analizado fue de 0.18 % en el suelo no contaminado, valor considerado como alto tomando en consideración lo indicado por la NOM-RECNAT-021-2000 (Cuadro 5).

6.1.7. Determinación de la Capacidad de intercambio catiónico y bases intercambiables

La capacidad de intercambio catiónico (CIC) es una propiedad química a partir de la

cual es posible inferir acerca del tipo de arcilla presente, de la magnitud de la reserva nutrimental y del grado de intemperismo de los suelos (NOM-RECNAT-021-2000). Los resultados obtenidos en este análisis se observan en el Cuadro 5. Con respecto a la reserva nutrimental se considera que ésta es abundante cuando la CIC es mayor a 25 $\text{cmol}_{(+)} \text{kg}^{-1}$ de suelo, para el caso de este suelo su CIC fue de 33.5 $\text{cmol}_{+} \text{kg}^{-1}$, que resultó ser alto. La CIC depende de la textura del suelo y del contenido de materia orgánica. En general, entre más arcilla y materia orgánica en el suelo, la capacidad de intercambio es mayor (Fernández *et al.*, 2006).

Debe tenerse en cuenta que el abastecimiento de nutrientes es la capacidad de los suelos para retener los iones. Los cationes tales como el NH_4^+ , K^+ , Ca^{2+} , y Mg^{2+} son extraídos de la solución por los minerales arcillosos, los cuales debido a su carga eléctrica negativa, atraen a los iones cargados positivamente. El intercambio iónico está estrechamente relacionado con las transformaciones biológicas debido a sus efectos sobre la disponibilidad de nutrientes y la acidez del suelo (Alexander, 1990). Algunas de las sustancias inorgánicas que asimilan los microorganismos son aniónicas. Estas sustancias aniónicas están representados por bicarbonatos, nitratos, fosfatos, sulfatos y molibdato. Sin embargo, el intercambio aniónico no es apreciable en el suelo por lo que es de poca importancia desde el punto de vista biológico (Alexander, 1990).

Las arcillas intervienen también en la fijación de los microorganismos. Aunque la superficie de arcilla y superficie de la célula microbiana tienen cargas negativas, la superficie de las arcillas son neutralizados por la acumulación de iones de carga positiva, tales como K^+ , Na^+ , Ca^{2+} , Mg^{2+} , Fe^{3+} y Al^{3+} . En conjunto, estas cargas superficiales negativas y positivas son llamadas doble capa eléctrica. Del mismo modo, los microbios tienen una doble capa eléctrica con respecto a la superficie de la arcilla. Cuando las fuerzas de repulsión se reducen por la presencia de iones con carga positiva, fuerzas de atracción electrostática permite la fijación de células microbianas a la superficie de las arcillas. Aunque la importancia principal de la CIC es intercambiar iones con respecto a la solución del suelo y poner a disponibilidad los nutrientes, también interviene en la fijación de los microorganismos a las arcillas con la formación

de la doble capa, evitando en cierto modo el lavado de los microbios por el agua (Raina *et al.*, 2000).

Con respecto a los resultados de los análisis de bases intercambiables (K^+ , Na^+ , Ca^{2+} , Mg^{2+} , Fe^{3+}) se obtuvieron valores altos, con base en lo establecido en el cuadro de interpretación de la NOM-021-RECNAT-2000, un suelo que tiene una cantidad de bases de intercambio alta respecto de la CIC es un suelo saturado y fértil.

6.2. Estudio de la Población de Hongos Filamentosos

En esta sección se presentan los resultados de la cuantificación de las poblaciones de hongos en los dos tratamientos. Se inició el día 4 (96 h), haciendo recuentos en los días 8 (192 h), 16 (384 h) y 32 (768 h). Esta medición se realizó por el método de recuento en placa de células viables por dilución seriada (Raina *et al.*, 2000).

El análisis estadístico (ANOVA) y las pruebas de comparación de medias poblacionales de los hongos nativos para el suelo contaminado y el suelo no contaminado, evidenciaron que para el día cuatro no hubo diferencia estadística significativa ($p=0.05$), es decir que los dos tratamientos tuvieron el mismo efecto en el tamaño de la población de hongos nativos cuantificados en unidades formadoras de colonias por gramo de suelo seco ($UFC\ g^{-1}\ ss$), siendo estas de $2.35E+04$ y $3.39E+04$ respectivamente (Figura 21).

Los resultados obtenidos al día 8 de evaluación, mostró que la población de hongos que mayor crecimiento presentó fue el suelo no contaminado con $2.16E+04\ UFC\ g^{-1}\ ss$ teniendo una significancia estadística alta ($p<0.05$), considerando la comparación de medias para este día de evaluación, los dos tratamientos presentaron medias poblacionales diferentes como se aprecia en la Figura 21.

Los resultados obtenidos los días 16 y 32 después de establecido el experimento, y de acuerdo al análisis de varianza, se encontró que no existió diferencia estadística

significativa ($p < 0.05$); con la prueba de comparación de medias de Tukey se observó que en los tratamientos evaluados, las UFC g^{-1} ss de hongos nativos fueron similares.

El tamaño de las poblaciones nativas de hongos totales se aprecia más reducido en el suelo contaminado, esto es evidente ya que en los cuatro tiempos analizados la biomasa fue relativamente menor en comparación al suelo no contaminado. Para el caso de la disminución progresiva de la población en el suelo no contaminado como se muestra en la Figura 21 se debe a la baja disponibilidad de nutrientes en el suelo que permitan la actividad metabólica del microorganismo. Manahan (2007), menciona que los microorganismos nativos pueden estimularse añadiendo nutrientes y aceptores de electrones para mejorar la degradación de los compuestos y formar biomasa. En este ensayo a los tratamientos no se les suministró nutrientes a los microorganismos.

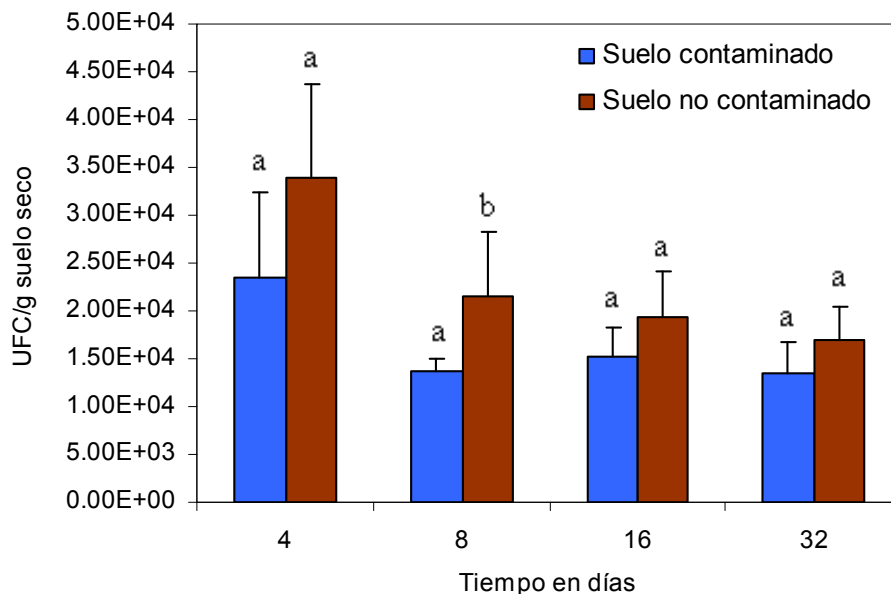


Figura 21. Cuantificación de la población de hongos nativos en los tratamientos evaluados. Letras distintas en cada par de barras indican diferencias estadística significativas (Tukey, $p = 0.05$).

Por las propiedades tóxicas del petróleo crudo las comunidades de microorganismos del suelo son alteradas por este compuesto, desapareciendo o disminuyendo las especies menos resistentes. Esta toxicidad del petróleo aumenta con la concentración de compuestos aromáticos de baja ebullición. La inhibición del crecimiento puede ser

debido a la intoxicación de la flora microbiana nativa por contacto directo con petróleo, especialmente petróleo crudo fresco. LaGreca *et al.* (1996), mencionan que cuando la concentración de contaminante es alta puede ocurrir que se produzca inhibición del crecimiento microbiano o intoxique a los microorganismos y estos mueran cuando la concentración es muy alta. Ercoli *et al.* (1998), han registrado en ensayos de campo que la actividad biológica de los microorganismos se ha llevado a cabo en suelo con concentraciones de hasta 30 % del contaminante. Observaron también que por encima de este valor la actividad microbiana es muy baja pero no nula.

6.3. Análisis Químico de Hidrocarburos

6.3.1. Análisis mediante el método de referencia

Para extraer los hidrocarburos del suelo se utilizó el método de reflujo con equipo Soxhlet, tomando como referencia el método 3540C de la USEPA (1996), en el cual la cuantificación se realizó por gravimetría. Ésta técnica mide el peso de los contaminantes totales extraídos con un solvente por medio de una balanza analítica. El método cuantifica los hidrocarburos que son extraídos por reflujo dentro de un solvente adecuado. La extracción por reflujo o Soxhlet es una de las técnicas analíticas más ampliamente usadas; es un procedimiento para la extracción de compuestos orgánicos no volátiles y semivolátiles de sólidos, como suelos, lodos y residuos. Este método asegura el contacto íntimo de la matriz de la muestra con el solvente de extracción. Para la óptima extracción de los compuestos orgánicos, los sólidos deben estar en partículas pequeñas; razón por la cual el suelo fue disgregado y pasado por un tamiz de dos milímetros de diámetro de malla (Fernández *et al.*, 2006).

Como resultados del proceso de extracción, aparte de los hidrocarburos totales de petróleo, se obtuvieron también los hidrocarburos biogénicos del tratamiento no contaminado. En el Cuadro 6 se presentan los resultados de los HTP para cada una de las repeticiones en cada tiempo analizado del suelo Gleysol, se indican también los valores promedios, las desviaciones estándar y los coeficientes de variación. Con

respecto a las concentraciones de HTP se observó que los coeficientes de variación (CV) de los tiempos 2, 4, 8 y 16 fueron inferiores al 5 % en todas las repeticiones; excepto para los tiempos 1 y 32 que mostraron un CV mayor a 5 %. El valor del CV indica la precisión de la cuantificación de los HTP durante el proceso de extracción por el método de referencia Soxhlet. Infante (2007) considera que en áreas de investigación donde se tienen datos de experimentos previos, el CV es muy usado para evaluar la precisión de un experimento; Reyes (1985) por su parte menciona que un CV igual al 5 % indica una variabilidad relativamente uniforme. Para el caso del presente trabajo se tiene que existe precisión en las determinaciones de HTP para cada tiempo evaluado en la experimentación.

En la Figura 22 se grafican las concentraciones de petróleo crudo y biogénico por tiempo, con respecto a los TPH se observa que las concentraciones para cada tiempo fueron disminuyendo, en el día uno de experimentación la cantidad obtenida de petróleo crudo fue de 29849 mg kg^{-1} que es contrastante con el resto de los tiempos evaluados, excepto para el día 32 en la que se observa una reducción importante, registrando una concentración de 18835 mg kg^{-1} de HTP. Es evidente que después de un día de establecido el experimento se presentó la mayor pérdida del hidrocarburo, para los tiempos restantes aunque la concentración siempre se fue abatiendo, ésta no fue tan marcada como en los días uno y dos.

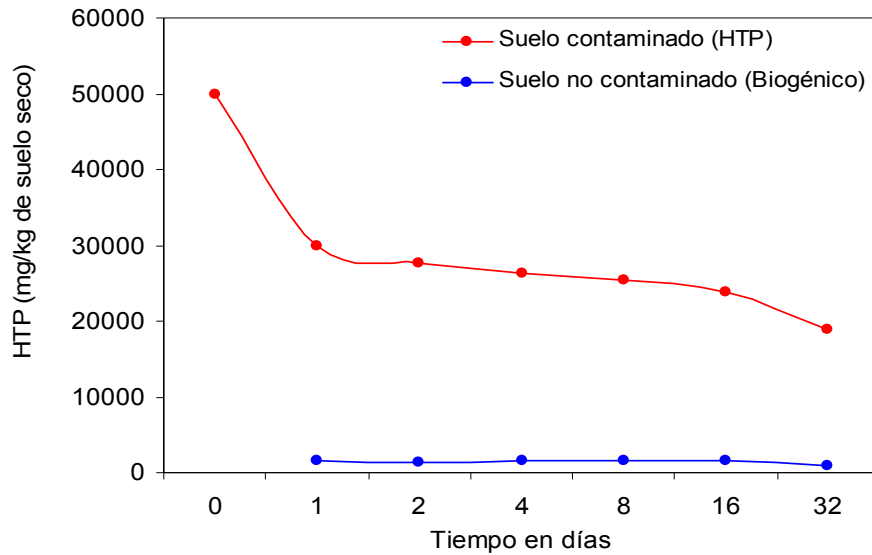


Figura 22. Concentración de petróleo crudo e hidrocarburos biogénicos por tiempo evaluado en suelo Gleysol. Los valores corresponden al promedio de seis repeticiones.

Para el caso de los hidrocarburos biogénicos, se observa que las concentraciones fueron relativamente constantes durante el periodo experimental. Los hidrocarburos de tipo biogénicos son sintetizados por casi todas las plantas, animales terrestres y marinos, incluyendo los microorganismos. Este tipo de hidrocarburos es el que se genera naturalmente y fue cuantificado en el suelo no contaminado con petróleo crudo. Aun cuando casi todas las plantas y animales terrestres pueden sintetizar hidrocarburos denominados como biogénicos, éstos exhiben un alto nivel de n-alcanos de número impar que pueden ser degradados durante la descomposición de la materia orgánica del suelo, de estos compuestos los de tipo aromático se forman por condensación durante la mineralización y dar lugar a sustancias de mayor peso molecular, como las que forman parte del humus (Bedair y Al-Saad, 1992). En temas de investigación los n-alcanos y sus índices han sido utilizados para distinguir entre aportes biogénicos y derivados del petróleo (Pelletier *et al.*, 1991). Las cadenas de carbono impar tienden a prevalecer en los materiales biológicos, en tanto que en los derivados petrolíferos hay una preferencia por los carbonos pares (Snedaker *et al.*, 1995).

En lo que se refiere al porcentaje de degradación del petróleo crudo en el Gleysol con respecto a su flora microbiana nativa, específicamente hongos filamentosos, evaluados

durante 32 días en condiciones controladas de temperatura y humedad; se observó que existe diferencias estadísticas significativas ($p=0.05$). Es decir, que existe un efecto altamente significativo de la variable tiempo ($p > F= 0.0001$) sobre la variable degradación de HTP. En la comparación de los promedios de degradación a tiempos diferentes mediante la prueba de Tukey y con un nivel de significancia de 5 %, se pudo observar que la degradación en el tiempo uno (20819 mg kg^{-1}), resultó ser altamente significativa en comparación al resto de las medias, es decir en el tiempo uno se presentó la mayor pérdida de petróleo crudo del suelo comparado con el resto de los tiempos analizados (Cuadro 7 y Figura 23).

En el Cuadro 7, se muestra que la mayor pérdida de petróleo crudo se dio en el día uno de experimentación con el 41.9 % (20819 mg kg^{-1} de HTP) y para el día 32 con el 9.9 % (4936 mg kg^{-1} de HTP). El incremento para el día 32 se puede deber a que los microorganismos después de un período de adaptación, posiblemente estarían en condiciones de iniciar la metabolización del petróleo crudo, considerando que el ensayo fue de 32 días; tiempo que se puede tomar como periodo de adaptación. De acuerdo con Levin y Gealt (1997), para que un microorganismo pueda conseguir una ruta degradadora de un contaminante, éste debe pasar por un periodo de adaptación suficiente en presencia de agentes contaminantes. Este incremento se presentó de forma significativa en las seis repeticiones para el día 32 del suelo contaminado, en comparación a las evaluaciones anteriores que fueron menores. La mínima pérdida se presentó en el día ocho de experimentación con el 1.5 % ($732.521 \text{ mg kg}^{-1}$ de HTP), para el resto de las evaluaciones la degradación se manifestó relativamente constante (3.3, 2.8 y 3.3 para los días dos, cuatro, ocho y 16 respectivamente). En la Figura 23 se presenta de manera gráfica, el porcentaje de degradación por tiempo evaluado. Aún cuando se considera la posible capacidad metabólica de los microorganismos nativos para la degradación de estos hidrocarburos, el porcentaje en pérdida de los TPH en el experimento, también debe ser atribuida a factores como los procesos de cambios físicos; entre ellos están: la fotólisis y la volatilización, esta situación normalmente implica a contaminantes orgánicos en las zonas saturada y no saturada del suelo. Por lo tanto, un compuesto químico presente en el suelo se volatiliza a la atmósfera, en este

proceso la presión de vapor es uno de los factores más importantes que gobierna la volatilización (Garrett *et al.*, 2000). Los componentes del petróleo muestran una absorción eficiente en un rango amplio del espectro del infrarrojo. Una película de petróleo crudo, absorbe una parte considerable de la energía, lo que lleva a la transformación y la degradación de la película a través de la oxidación fotoinducida (Riser-Robert, 1992). Por otra parte, la absorción de luz en el rango de la radiación visible y ultravioleta, puede modificar, también, la energía interna de las moléculas de hidrocarburo propiciando transiciones electrónicas entre niveles energéticos, proceso conocido como fotoquímico (Doménech, 1999). Otros procesos se relacionan con los cambios químicos como la hidrólisis, reducción y oxidación, estos dos últimos son los más importantes porque pueden transformar los contaminantes orgánicos. Tanto los procesos fotoquímicos como los químicos experimentan una transferencia electrónica que involucra a compuestos orgánicos en el sistema agua-arcilla (Garrett *et al.*, 2000; Ney, 1990; Fernández *et al.*, 2003).

Considerando los resultados anteriores se puede asumir que la mayor pérdida de los hidrocarburos del petróleo en las primeras 24 h se debe principalmente a fenómenos físicos o químicos como la volatilización de las fracciones más ligeras del crudo, y al hecho de que en el presente trabajo se utilizó un petróleo crudo que de acuerdo a sus grados API es un hidrocarburo superligero, contiene por lo tanto más fracciones de estructura química sencilla y bajo peso molecular (C_5 a C_{10}) siendo más volátiles, de tal manera que con temperaturas relativamente altas estos compuestos se evaporan (IMP, 2008).

Debido al alto porcentaje de remoción, sobre todo en el tiempo uno, este no puede ser atribuido totalmente a los microorganismos nativos, especialmente a los hongos. Los bajos porcentajes de remoción en los siguientes tiempos analizados, tal vez se deba a la limitada capacidad metabólica de estos microorganismos, sin embargo la volatilización puede seguir siendo la causa principal, considerando que las fracciones ligeras de hidrocarburos tienen un menor punto de ebullición (0-20 °C), estando constituidos por fracciones gaseosas y líquidas. Viñas (2005), menciona que las

parafinas volátiles, *n*-alcanos e isoprenoides de un tamaño C₁-C₁₀, es la fracción más volátil del crudo y por lo tanto la más susceptible de pérdidas abióticas por volatilización. Considerando que el tratamiento contaminado con petróleo crudo "Olmeca" denominado superligero, se mantuvo siempre por encima de los 30 °C, este factor tuvo influencia de forma determinante en el proceso de volatilización.

Cuadro 6. Resultado de las determinaciones de Hidrocarburos Totales de Petróleo (HTP) por tiempos en el suelo Gleysol.

Tiempo (días)	Concentración de TPH por repetición (mg kg ⁻¹ suelo seco)						Media	Desviación estándar	C. V (%)
	1	2	3	4	5	6			
1	28919	29616	28571	33333	28803	29849	29849.0	1776.6	5.95
2	28339	27642	28339	27758	26829	26364	27545.5	802.7	2.91
4	26248	26521	27293	26015	25216	25481	26129.3	747.1	2.86
8	26016	25551	24622	25900	25087	25203	25396.8	527.9	2.08
16	24041	23693	22532	24738	24854	22764	23770.8	973.5	4.10
32	20905	18118	19860	17886	18350	17886	18834.6	1256.1	6.67

Cuadro 7. Degradación del petróleo crudo ligero tipo Olmecca en Gleysol con hongos nativos.

Tiempo de evaluación (días)	Degradación (mg kg ⁻¹ suelo)	Degradación (%)
0	50000**	0
1	20818 ^{a*}	41.9
2	1635*	3.3
4	1416 ^{c*}	2.8
8	732 ^{c*}	1.5
16	1626 ^{c*}	3.3
32	4936 ^{b*}	9.9

*Promedio de seis repeticiones. Letras distintas son estadísticamente diferentes (Tukey, p=0.05).

**Concentración inicial.

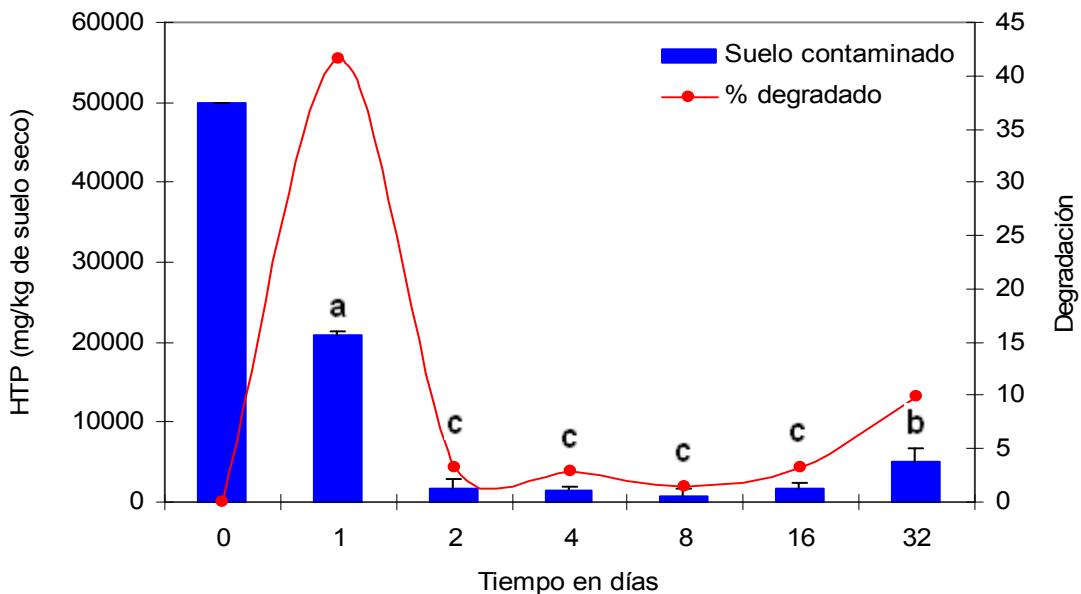


Figura 23. Pérdida de HTP por tiempo en el tratamiento con suelo contaminado uno. Letras distintas indican diferencias estadística (Tukey, p=0.05).

6.4. Predicción de la Degradación del Petróleo Crudo en el Suelo Gleysol

Para el desarrollo del modelo de predicción se llevó a cabo una regresión lineal (modelo trigonométrico de primer grado), que fue ajustada utilizando el procedimiento PROC REG de SAS, cuyo modelo viene dado por:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

El análisis de regresión es una técnica para investigar y modelar la relación entre variables. El uso de la ecuación de regresión resulta útil cuando se emplea para predecir el valor de una variable, a partir del valor de otra variable. Estos modelos son usados con varios fines, que incluyen descripción de datos, estimación de parámetros, así como para realizar predicciones y estimaciones (Montgomery *et al.*, 2005; Triola, 2006).

Con base en los resultados de $p > F = 0.0029$, indica que el modelo es significativo. El coeficiente de determinación (r^2) cuyo valor es 0.9128, indica el porcentaje de variabilidad de los datos experimentales que se explica con el modelo, en este caso el modelo explica en 91.2 % la variación de la degradación del petróleo crudo con respecto al tiempo. El grado de precisión con que se compararon los tiempos analizados está dado por el coeficiente de variación (CV) que resultó ser de 4.92 %, que es un buen índice de la confiabilidad del análisis. En general, cuanto más pequeño es el CV mayor confiabilidad hay en los datos colectados (Reyes, 1985), para este caso se considera que la variación fue menor al 5 % por lo que los datos fueron reproducibles.

Finalmente, se obtuvo el modelo estimado en el que el proceso de degradación se ajusta a la ecuación $HTP(\text{mg kg}^{-1}) = 32039 - 1925.6(\text{tiempo})$. Con base en el grado de significancia (confiabilidad de un 95%), r^2 y el CV, se puede concluir que el modelo matemático obtenido, es suficiente para predecir la degradación del petróleo crudo en función del tiempo para este tipo de suelo y en las mismas condiciones experimentales (Figura 24).

Para su aplicación en campo (áreas contaminadas por petróleo), el modelo debe tener ajustes por humedad y temperatura.

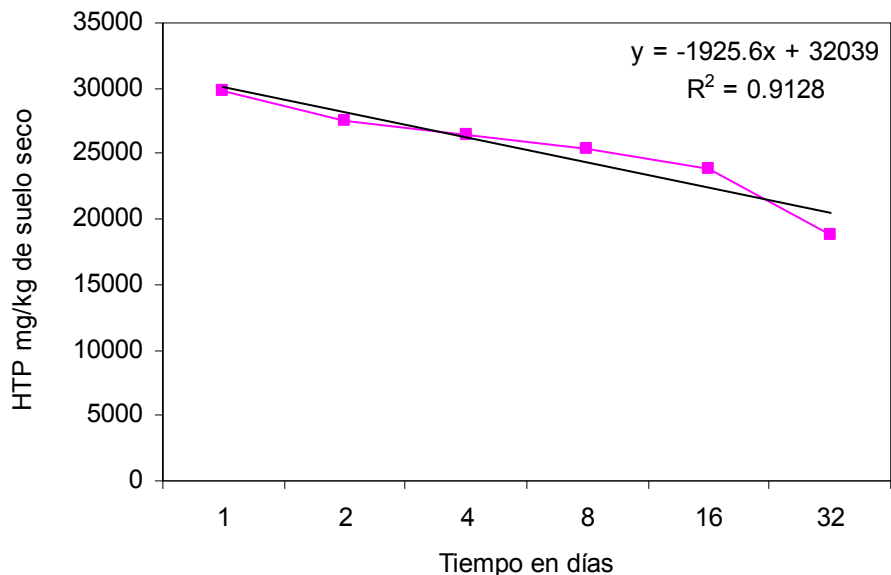


Figura 24. Línea de regresión y modelo de predicción para la degradación de HTP y tiempo del bioensayo.

6.5. Interpretación de los Espectros NIR

El espectro de las muestras de suelo del experimento, se registró en la modalidad de reflectancia en celda de cuarzo, realizando lecturas en ocho tiempos distintos para la generación de los espectros de los dos tratamientos, contaminados y no contaminados con petróleo crudo. Por cada tratamiento se realizaron seis repeticiones con igual número de registros en el instrumento NIRS.

En el caso de los datos obtenidos en las lecturas de las 12 UE y dado que el equipo genera un dato cada dos nm en un rango que va de los 400 hasta los 2500 nm; se generaron en total de 1050 variables por muestra; a partir de estos datos se genera una matriz formada por 88200 datos o variables. Las 1050 variables al ser graficadas representan la huella espectral característica de cada muestra como es mostrada en la Figura 25.

En la Figura 25 se observa también que los dos tratamientos, suelo contaminado (color rojo) y no contaminado (color verde) se diferencian notablemente en las primeras lecturas realizadas en NIR, conforme avanzó el experimento esta diferencia fue

disminuyendo. Es decir, que la separación entre la muestra contaminada y no contaminada cada vez fue menor, de tal forma que los espectros de los dos tratamientos en la última lectura en NIR se acercan entre si como se aprecia en la Figura 27.

El espectro generado por el instrumento NIR presenta tres regiones, la primera región UV-Vis (400-700 nm), la región de transición UV-Vis/NIR (702-1098 nm) y la región NIR (1100-2500 nm) (WinISIS III, 2005) (Figura 26).

La región UV-Vis (400-700 nm) que es la región visible, en ella se absorbe la energía lumínica y es reflejada como vibración electrónica (recuadro verde, Figura 26).

La segunda región (recuadro azul) corresponde a la transición entre UV-Vis y NIR (702-1098 nm), en esta región se identifican algunos componentes presentes en la muestra tales como carbohidratos y celulosa que se encuentran en las regiones de 858 a 868, 876 a 884, 894 a 906 y 916 a 918 nm; proteínas en las regiones 870 a 874, 908 a 910, 1014 a 1032 y 1048 a 1052 nm; aceites en las regiones 912 a 914, 930 a 932, 962 a 968, 998 a 1002 y 1008 a 1012 nm; agua en las regiones 934 a 960, 984 a 996 nm; y azúcares en la región 1004 a 1006nm. En la Figura 26 se pueden observar las bandas de absorción correspondientes a los compuestos característicos del suelo Gleysol (agua, proteína, aceite, carbohidratos, etc.) (WinISIS III, 2005).

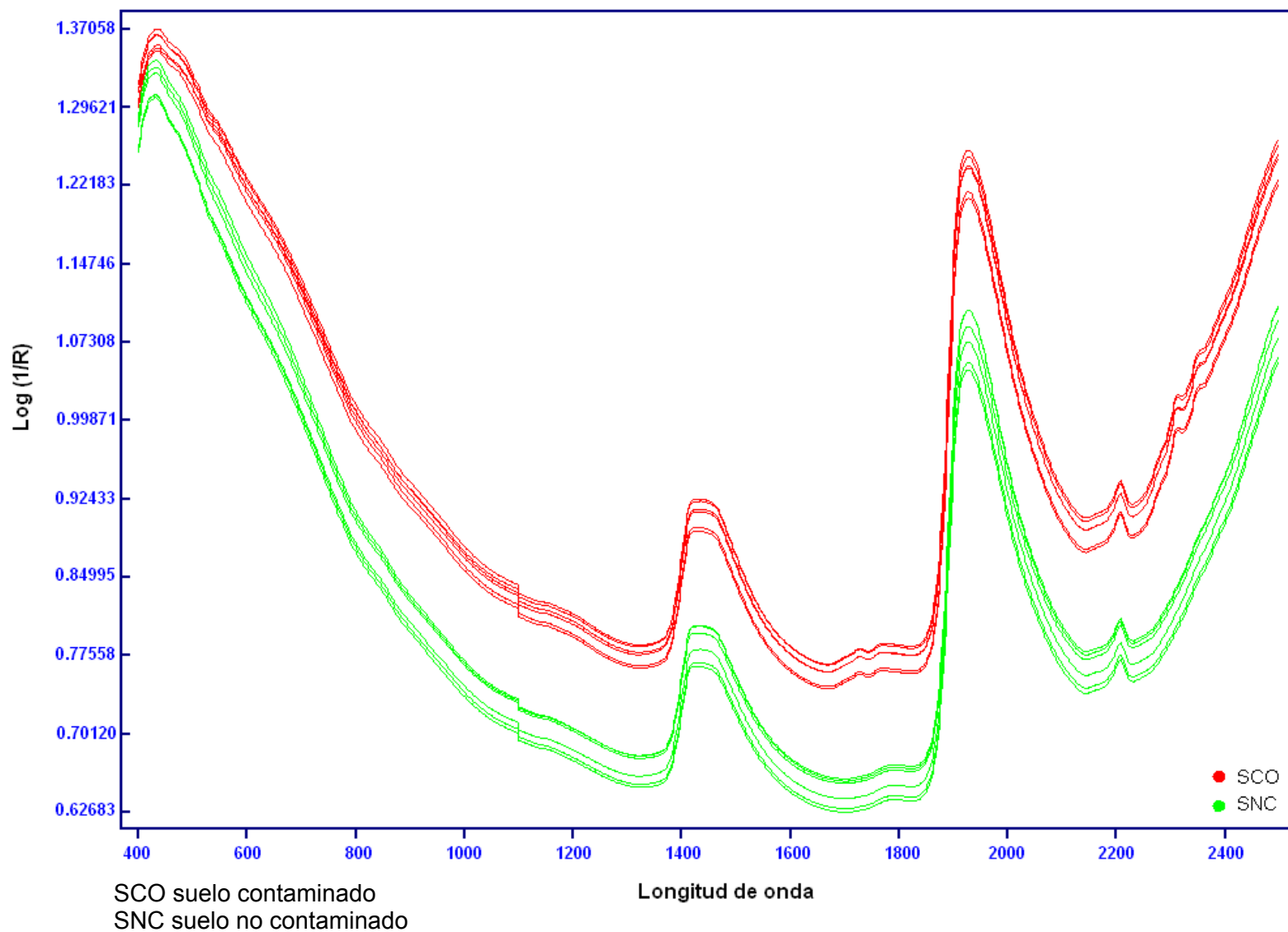


Figura 25. Espectros NIR característicos de las muestras de suelo no contaminado y contaminado con HTP, registrados al día uno de iniciado el experimento.

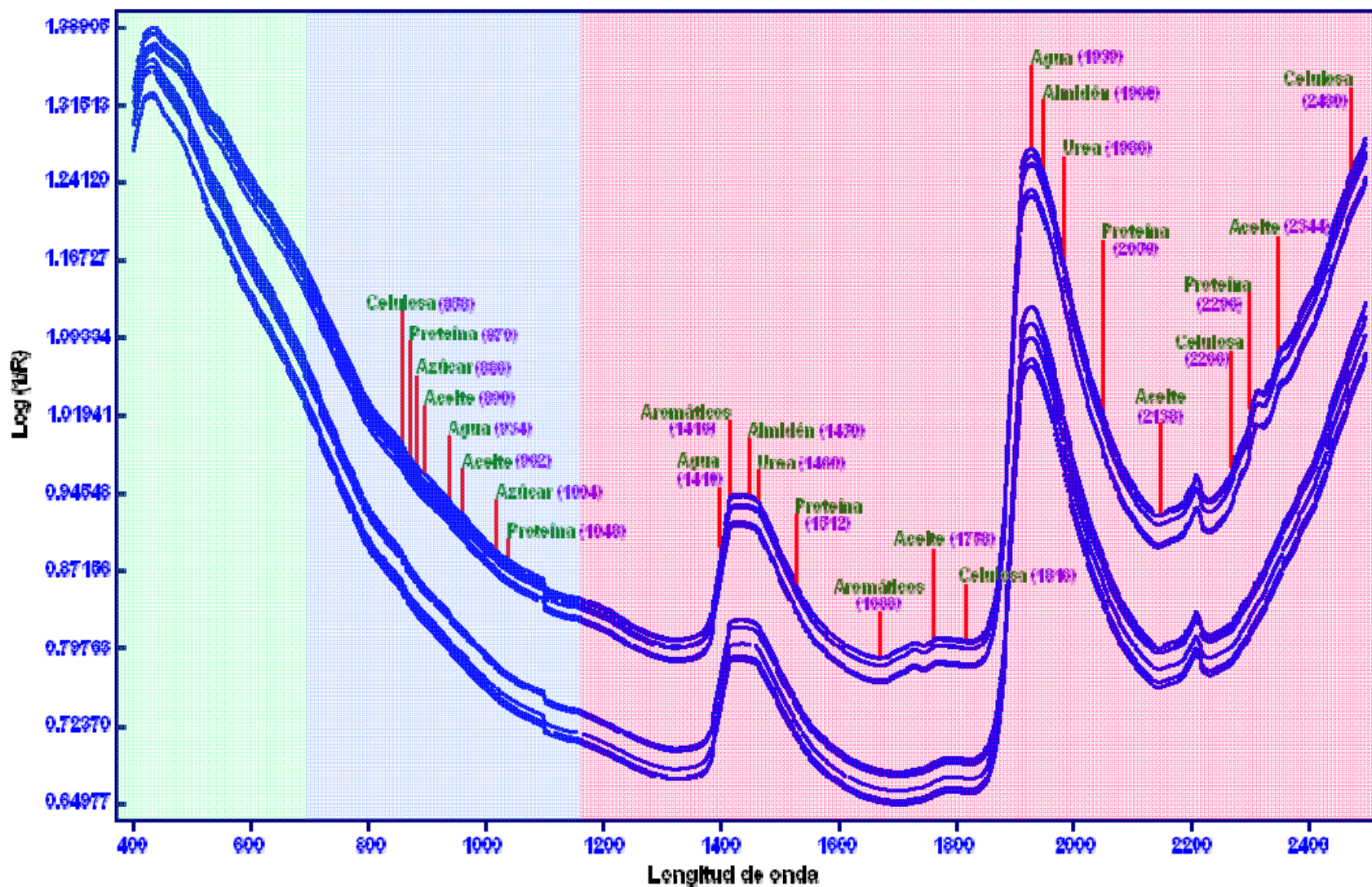


Figura 26. Interpretación de los espectros NIR de las muestras de suelo no contaminados y contaminados con HTP (WinISIS III, 2005).

La tercera región (recuadro rojo) corresponde a la región espectral NIR (1100–2500 nm) en donde se presenta únicamente absorción de energía reflejada en estiramiento y doblez de las moléculas, aquí se pueden identificar algunos compuestos como el agua que se observaron en las regiones 1406 a 1408, 1786 a 1796 y 1936 a 1946 nm; compuestos aromáticos en las regiones 1416 a 1418, 1444 a 1418 y 1680 a 1690 nm; aceites en las regiones 1410 a 1416, 1758 a 1762, 2066 a 2076, 2136 a 2146, 2304 a 2308, 2344 a 2348 y 2376 a 2386 nm; carbohidratos en las regiones 1444 a 1448, 1452 a 1460, 1536 a 1546, 1920 a 1930, 1956 a 1966, 2096 a 2104, 2276 a 2286 y 2320 a 2336 nm; urea en las regiones 1460, 1484 a 1496, 1516 a 1526, 1986 a 1996, 2026 a 2036 y 2066 a 2076 nm; celulosa en las regiones 1484 a 1496, 1770 a 1786, 1816 a 1826, 2096 a 2104, 2266 a 2276, 2330 a 2340, 2350 a 2354 y 2480 a 2490 nm; y proteínas en las regiones 1506 a 1516, 2050 a 2064, 2176 a 2186, 2296 a 2306 y 2466 a 2476 nm (Figura 26) (WinISIS III, 2005).

Sobre el conjunto de espectros obtenidos se llevó a cabo un pretratamiento de datos correspondiente al análisis de componentes principales (PCA), finalmente se realizó un análisis discriminante lineal (LDA), un modelado suave independiente de analogía de clases (SIMCA) y una regresión por mínimos cuadrados parciales (PLS) como análisis clasificadorio de los datos.

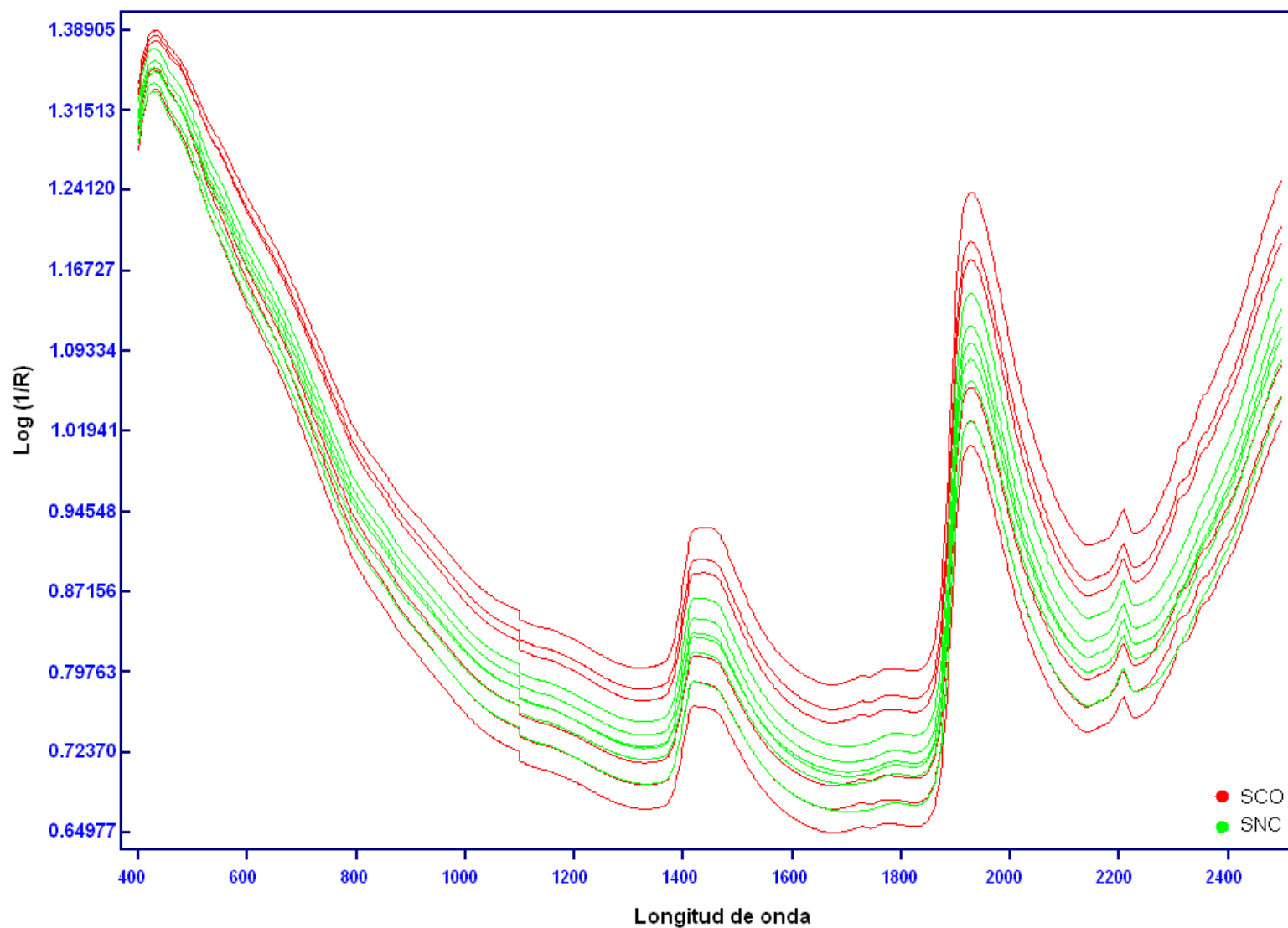


Figura 27. Espectros NIR de las muestras de suelo no contaminado y contaminado con HTP, a los 32 días de experimentación.

6.5.1. Pretratamiento de los datos espectrales

En el pretratamiento espectral se han incluido tanto las transformaciones de variables como las técnicas de reducción de dimensiones, todas ellas encaminadas a eliminar el ruido, para mejorar la calidad de la señal y hacer a la matriz de datos más apropiada para los tratamientos estadísticos posteriores (Massart *et al.*, 2003). Los pretratamientos espectrales son correcciones del espectro basadas en tratamientos matemáticos con la finalidad de aumentar la calidad de la señal instrumental, reduciendo al máximo la información no asociada a la propiedad estudiada (Bautista, 2009).

A los datos espectrales obtenidos se les realizó un pretratamiento, con la finalidad de reducir efectos o componentes. Estos efectos no deseados presentes en las señales obtenidas en espectroscopia se denominan comúnmente ruido. El ruido es atribuido a los componentes de la instrumentación utilizada para el registro del espectro (ruido instrumental), a las variaciones de temperatura, humedad u otras condiciones ambientales durante el análisis (ruido ambiental) o bien variaciones en la señal debidas a la propia naturaleza de la muestra (Williams, 1987).

Con la finalidad de reducir o eliminar la contribución del ruido, a los datos espectrales se les realizó los tratamientos por Variable Normal Estándar (SNV) y primera derivada. La SNV utilizado como pretratamiento es con la finalidad de corregir efectos de tamaño de partícula en espectroscopia NIR (Barnes y Dhanoa, 1989). En este caso, la SNV se aplica individualmente a cada espectro, sin depender de ningún espectro de referencia. De esta forma se obtiene una escala común a todos los espectros. El procedimiento de corrección consiste en centrar cada espectro, restando el valor medio de absorbancia del espectro a cada uno de los valores de absorbancia a cada longitud de onda, dividiendo el resultado por la desviación estándar del espectro (Barnes y Dhanoa, 1989).

El uso de las derivadas es uno de los pretratamiento más utilizado en espectroscopia para minimizar desviaciones de línea base causada por efectos de dispersión. En

espectroscopia NIR se utiliza para disminuir los problemas más característicos: solapamiento de bandas y variaciones de línea base. La utilización de la primera derivada elimina los términos constantes a todas las longitudes de onda, es decir, desplazamientos de línea base, mientras que la segunda derivada corrige además las desviaciones causadas por los términos que varían linealmente con la longitud de onda. La aplicación de las derivadas permite un aumento de la resolución de bandas, pero por contra aumenta el ruido (Norris y Williams, 1984; Savitzky y Golay, 1964).

Para el presente trabajo fue necesario aplicar más de un pretratamiento para la corrección de los datos espectrales obtenidos de las muestras de suelo contaminado y no contaminado que permitieron eliminar en mayor medida el ruido espectral y obtener la información útil para establecer los modelos de clasificación y de estimación. De acuerdo con Dhanoa *et al.* (1994), en ocasiones resulta necesario utilizar más de un método de pretratamiento espectral para mejorar la calidad de los resultados.

6.5.1.1. Análisis de componentes principales (PCA)

Previo al tratamiento quimiométrico para la creación de las reglas de clasificación y el modelado de clases, se aplicaron técnicas no supervisadas para visualizar el comportamiento de los datos generados por los espectros NIRS.

Con el Software S-PARVUS se realizó un análisis exploratorio mediante el análisis de componentes principales (PCA). El PCA es un modelo matemático para organizar la información en un conjunto de datos, principalmente cuando se está en presencia de un número grande de variables. La estrategia de PCA, es descubrir nuevas variables llamadas Componentes Principales (PC) que expliquen la máxima variabilidad de los datos de modo que los residuales sean los más pequeños posibles. Esto permite describir la información con una cantidad de variables considerablemente menor que la cantidad original (Nuñez *et al.*, 2009). El PCA se considera una técnica exploratoria y de pretratamiento de los datos (reducción de dimensiones), como paso previo a la obtención del modelo de calibración o de clasificación. El método usado para el análisis

de PCA es el algoritmo NIPALS (*Nonlinear Iterative Partial Least Squares*), que permite obtener sólo los primeros PC sin necesidad de calcular todos los vectores propios de una matriz.

En PCA se suele utilizar el concepto varianza explicada, que mide el porcentaje de la varianza total que es explicado hasta la componente principal en cuestión (se trata de una varianza acumulada) (Massart *et al.*, 1997; Romaniuk *et al.*, 2007). Si se consigue encontrar pocos componentes (k) capaces de explicar casi toda la varianza total, podrán sustituir a las variables originales. Normalmente, sólo las primeras PC contienen información útil, mientras que las últimas describen ruido. Por ello, sólo las primeras PC son útiles en lugar del conjunto de datos de partida, no sólo por reducir la complejidad de la matriz de datos, sino también porque asegura que el ruido no es confundido con información.

Considerando lo anterior, la técnica de análisis exploratorio aplicado a los datos espectrales NIR, fue el PCA que permitió reducir el número de variables obtenidas al momento de la lectura de las muestras. Es decir, que con esta técnica los 1050 datos o variables espectrales originales se redujeron para obtener nuevas variables, o PC, con lo cual con sólo 10 PC es explicado aproximadamente el 100 % de la varianza total de las muestras.

Con los componentes principales calculados a partir de las 1050 señales espectrales NIR, se encontró que la máxima variabilidad de las muestras queda expresada en los dos primeros PC, ya que el primer componente principal (PC1) nos explica el 65.05 % de la variabilidad de los datos y el segundo (PC2) un 14.97 % de la restante, es decir, que en conjunto ambos componentes principales nos expresan el 80.02 % de la variabilidad total de las muestras de suelo contaminado y no contaminado. La variabilidad total explicada en los dos primeros componentes principales es considerada aceptable, lo cual indica que entre estos dos componentes se presentan casi el total de la variabilidad que inicialmente se encontraba expresada en las 1050 variables originales.

Los primeros PC obtenidos a partir del PCA reflejan un valor cercano al cien por ciento del total de la varianza, el resto de los componentes presentan una aportación mínima. En la Figura 28 se observa el gráfico de datos (PC2 *versus* PC1) obtenido de un PCA de los espectros mediante las transformaciones SNV y primera derivada en el intervalo espectral completo de las muestras de suelo contaminado y no contaminado con petróleo crudo.

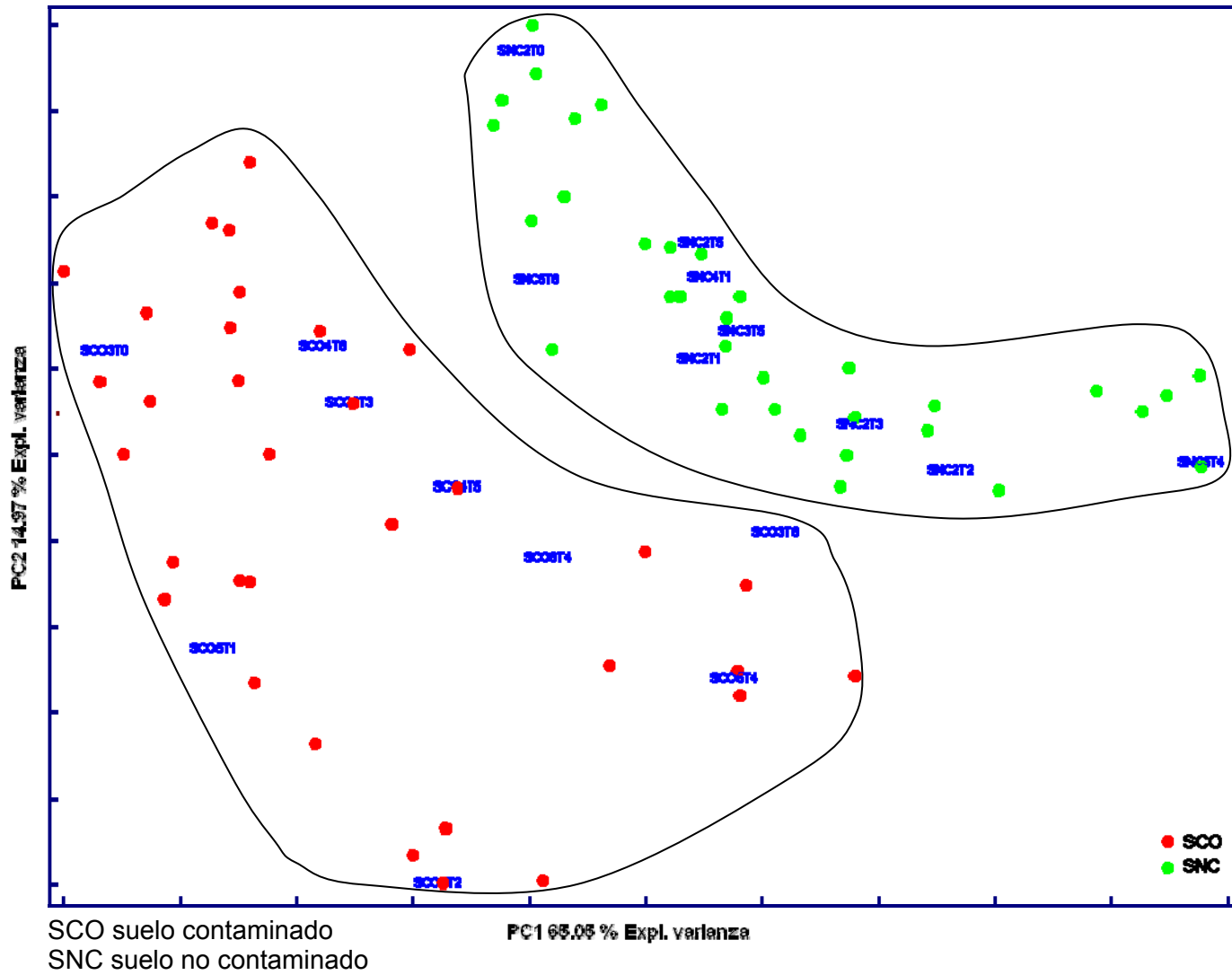


Figura 28. Proyección de las muestras de suelo en los dos primeros componentes principales.

La Figura 28 fue obtenida utilizando el primer y segundo componente principal, por ser éstos los de mayor poder discriminante (2.03698 y 0.73362, respectivamente) de acuerdo con la prueba de Fisher Weight para las categorías SCO y SNC (Cuadro 8).

Cuadro 8. Poder discriminante de los Componentes Principales (CP) para las categorías SCO y SNC.

Orden	Componente principal	Prueba de Fisher Weigt
1	PC1	2.03698*
2	PC2	0.73362*
3	PC4	0.24451
4	PC3	0.22807
5	PC6	0.00749
6	PC5	0.00473
7	PC8	0.00272
8	PC7	0.00098
9	PC10	0.00065
10	PC9	0.00001

* CP con mayor poder discriminante.

En la Figura 28 es presentada la distribución espacial de los objetos, en la que se observan dos grupos perfectamente definidos, estos grupos corresponden a las muestras de suelo no contaminado y contaminado con petróleo crudo. Para el caso del suelo contaminado su dispersión fue mayor con respecto al suelo no contaminado. En la Figura 28 se observa que las muestras pertenecientes al suelo no contaminado con petróleo crudo forman un grupo compacto y que sobre éste tiene más influencia el primer componente (PC1), mientras que para las muestras de suelo contaminado con petróleo crudo forman un grupo menos compacto, para el que se observa una variabilidad explicada a lo largo del componente principal dos (PC2).

Concluido el análisis exploratorio de PCA de las 12 muestras de suelos contaminados y no contaminados con petróleo crudo, caracterizados inicialmente por el método Soxhlet, se puede afirmar que la información generada con la tecnología NIRS es suficiente y útil para diferenciar los dos grupos identificados con los métodos convencionales. El siguiente paso fue generar modelos de clasificación mediante técnicas de reconocimiento de pautas supervisadas.

6.6. Análisis Clasificadorio

En el análisis clasificadorio se construyen modelos capaces de pronosticar la pertenencia de un objeto a una categoría sobre la base de las características del objeto. La matriz de datos contiene al menos una variable categórica, que indica la categoría a la que pertenece cada objeto y que constituye la respuesta o variable que se quiere predecir, y una o más variables de escala que describen otras tantas características de los objetos y que se utilizan como variables predictoras.

6.6.1. Análisis discriminante lineal (LDA)

Previo a la aplicación de esta técnica clasificatoria, a los datos se les realizó un pretratamiento de “autoscaling” y primera derivada, para finalmente aplicar la técnica de análisis discriminante lineal (LDA) sobre componentes principales, con el propósito de obtener el modelo de clasificación. La limitante de esta técnica, es que no es capaz de trabajar con matrices de datos demasiado grandes. La finalidad del LDA es establecer reglas de clasificación o asignación de unidades a los grupos, basadas en las puntuaciones de las variables independientes, a través del LDA se obtiene una proyección de los datos en un espacio de menor (o incluso igual) dimensión que los datos entrantes, con el fin de que la separación de las clases sea la mayor posible (García, 2009).

Para obtener el modelo se utilizó el 80 % de los datos como grupo de entrenamiento y el 20 % restante como grupo de validación, esta selección se realizó mediante el algoritmo de Kennard-Stone, de la herramienta KENNARD, del paquete V-PARVUS 2008. El modelo fue validado por el método de validación cruzada (*cross-validation*), con 5 grupos de cancelación.

Esta validación consiste en la utilización de los datos del conjunto de calibración para validar el modelo, aquí el conjunto de calibración inicial se divide en diferentes grupos de cancelación. Cada vez que se calcula el modelo se deja un grupo de cancelación

fuera y se calcula la ecuación de la regresión con los demás datos. Una vez realizada la calibración se prueba el modelo con los datos del grupo de cancelación excluido (Stone, 1974). Los resultados del LDA se muestran en el Cuadro 9.

Cuadro 9. Matrices LDA (de habilidad) de clasificación (a) y predicción (b) para las muestras de suelos contaminados y no contaminados con petróleo crudo.

(a).

Clase	Asignado a la clase	
	SCO*	SNC*
SCO	165	0
SNC	0	165
% correcto de clasificación	100	100
Habilidad de clasificación 100%		

(b).

Clase	Asignado a la clase	
	SCO	SNC
SCO	33	0
SNC	0	33
% correcto de predicción	100	100
Habilidad de predicción 100%		

* SCO suelo contaminado
SNC suelo no contaminado

Con base en el LDA se encontró que la habilidad de clasificación del modelo fue del 100%; mismo resultado para la habilidad de predicción que también fue de 100%. Esta última es la probabilidad de que cada una de las muestras este ubicada en su clase respectiva. Esto se puede apreciar mejor en la Figura 29 que se muestra a continuación la cual se generó utilizando la diferencia de puntuaciones, de la primera y segunda categoría.

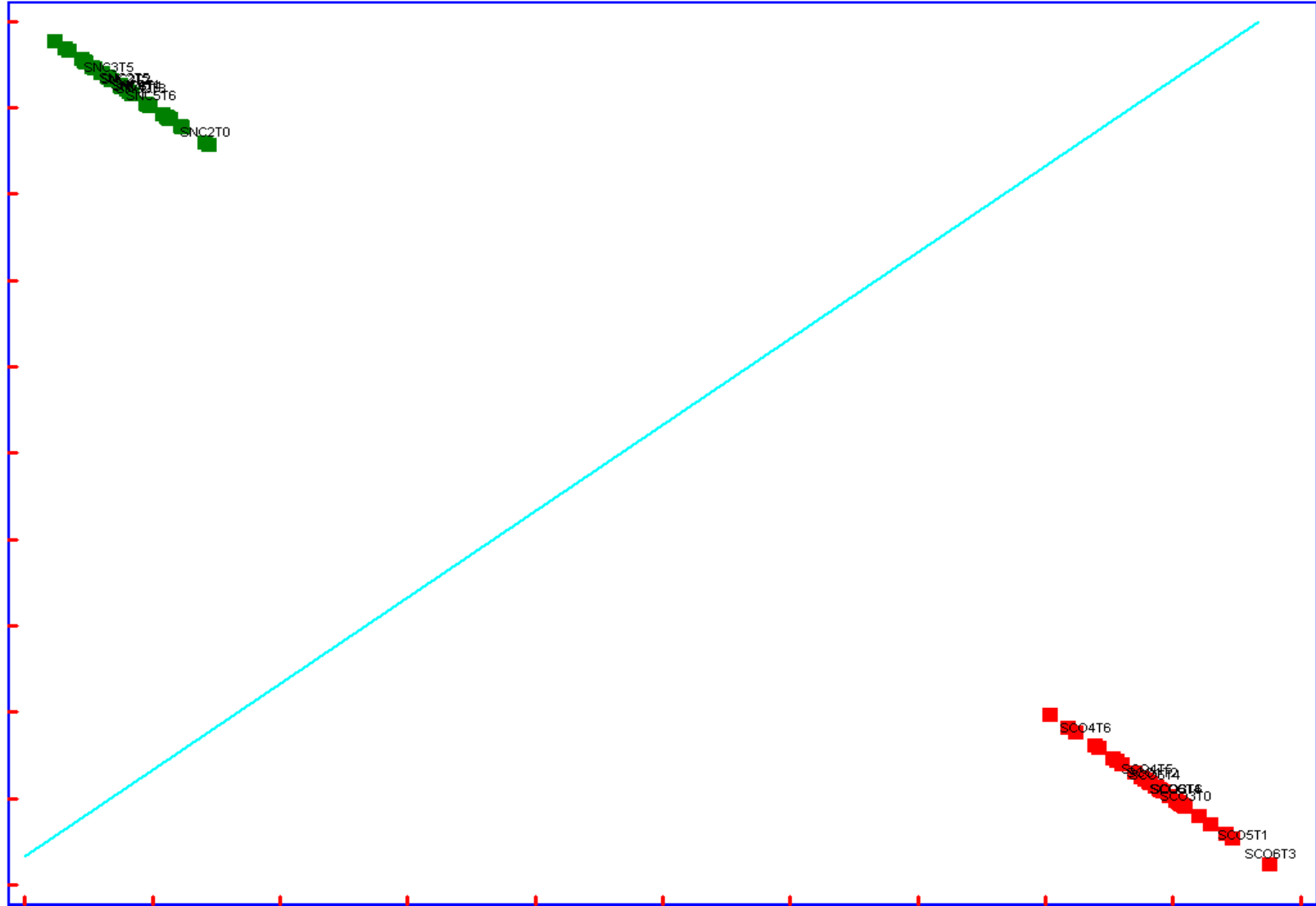


Figura 29. Modelo de clasificación LDA, obtenido con los espectros NIRS de las muestras de suelo contaminado y no contaminado con petróleo crudo. ■ SCO Suelo contaminado ■ SNC Suelo no contaminado.

La técnica quimiométrica LDA aplicada, resulta útil ya que las habilidades de clasificación y predicción en porcentaje fueron máximas, logrando con ella la completa separación y diferenciación en los dos grupos de muestras. En la Figura 29 se observa muy claramente los dos grupos de suelos totalmente separados, uno contaminado y el otro no contaminado con petróleo crudo, ubicados en sus respectivas clases. Se puede ver que los indicadores de grupos en las observaciones se disponen en dos grupos distintos.

6.6.2. Modelado suave independiente de analogía clases (SIMCA)

Con la finalidad de establecer un modelo para discriminar entre muestras de suelo contaminado y suelo no contaminado, los datos generados por NIR fueron analizados con la técnica de modelado suave independiente de analogía de clases (Soft Independent Modelling of Class Analogy). Esta técnica se basa en encontrar los ejes de los componentes en el espacio de las mediciones multivariantes para cada categoría y distribuye las muestras en cuatro clases: las clases A y B, muestras que pertenecen a ambas clases y muestras que no pertenecen a la clase A, ni a B. El método indica si la muestra analizada corresponde o no a un grupo de muestras de referencia indicando además a qué población pertenece dicha muestra (Beebe *et al.*, 1998).

Esta técnica se fundamenta en el análisis de componentes principales, mediante los cuales se construye un modelo independiente para cada una de las dos clases, previo a la aplicación de esta técnica, a los datos se les realizó un pretratamiento de “autoscaling” y primera derivada, con lo cual se pudo observar que las muestras son clasificadas en sus respectivas clases (Figura 30).

SIMCA se aplicó como una técnica de modelado sobre componentes principales en datos autoescalados, los modelos fueron creados utilizando el 80% de las muestras como grupo de entrenamiento y el 20% restante como grupo de prueba, el modelo fue validado mediante “*cross validation*” con 5 grupos de cancelación, con un 95% de confiabilidad, la selección de las muestras para la generación del grupo de

entrenamiento y el grupo de prueba se definió mediante el algoritmo Kennard-Stone, de la herramienta KENNARD del paquete V-PARVUS 2008, los resultados se muestran en el diagrama de Coomans (Figura 30), que permite realizar la comparación entre los dos modelos generados para sus respectivas clases.

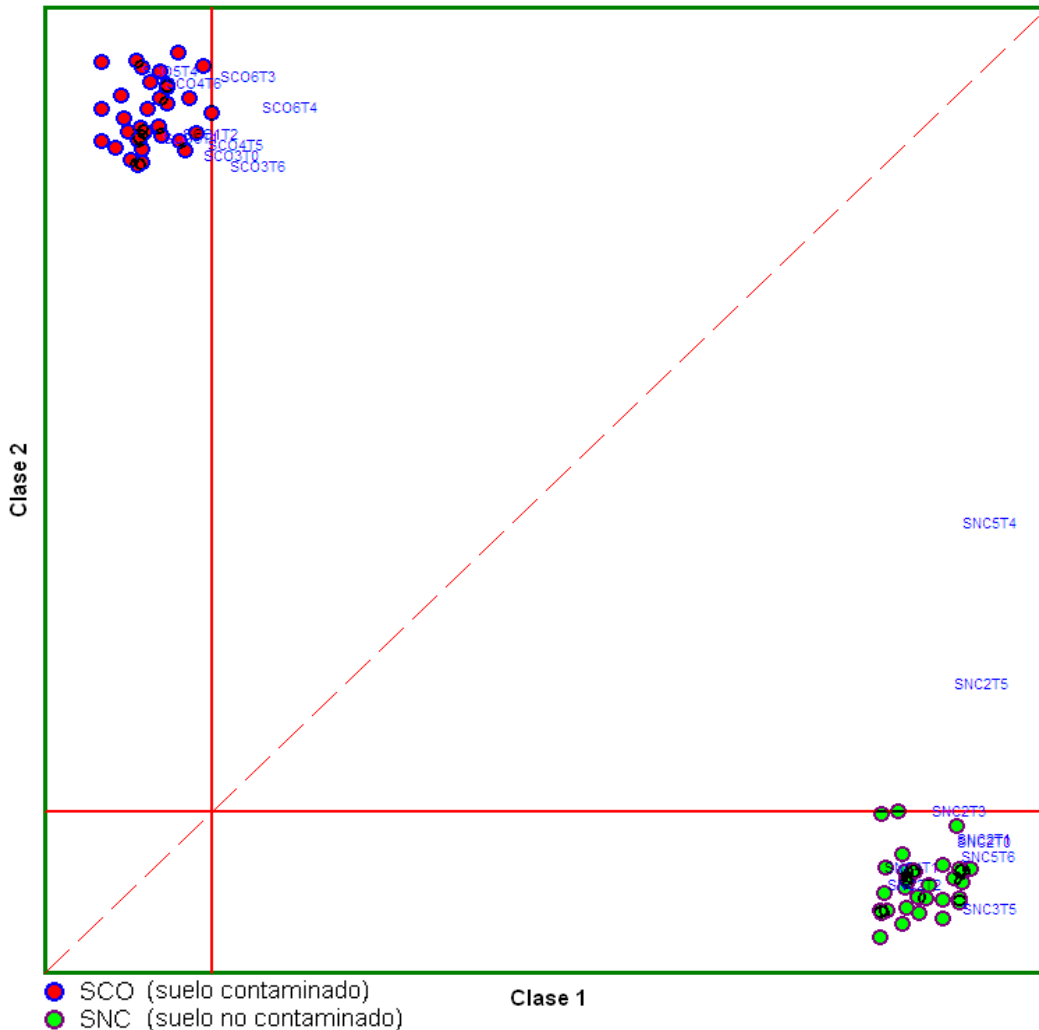


Figura 30. Diagrama de Coomans (SIMCA) obtenidos con los datos espectrales de NIR de suelo Gleysol no contaminado y contaminado con petróleo crudo.

En este diagrama se muestra que al comparar la clase uno, correspondiente al suelo no contaminado (SNC), con la clase dos propio del suelo contaminado (SCO); cada categoría se ubica en su respectivo grupo de acoplamiento del modelo. En el diagrama se observa que en el cuadro inferior izquierdo común a ambas clases, no existen muestras, lo que indica la especificidad de los modelos, para el caso del cuadro

superior derecho o zona de no pertenencia a los grupos A ni a B, tampoco se observa la presencia de muestras. Este resultado confirma la eficiencia del modelo al separar completamente a los dos suelos en base a la presencia del contaminante en ellos, indica también que la sensibilidad del modelo fue del 100%; ya que todas las muestras que componen el modelo son incluidas en el área representativa de éste, así mismo la capacidad de predicción y clasificación del modelo resultó ser del 100%.

Los objetos resultantes de esta etapa aportan un alto grado de información, tal es el caso de la predicción de clases, los errores de clasificación, el poder de discriminación de las variables y el poder total de modelación (Wold y Sjöström, 1977).

Al ensayar el modelo con el grupo de validación, cada una de las muestras desconocidas para los modelos, fue clasificado y predicho correctamente, los resultados son presentados en el Cuadro 10.

Cuadro 10. Resultados SIMCA. Porcentaje de correcta clasificación (a) y predicción (b) para el suelo Gleysol contaminado y no contaminado.

(a).

Clase verdadera	Asignado a la clase	
	SCO	SNC
SCO	33	0
SNC	0	33
% correcto de clasificación	100	100
Habilidad de clasificación 100%		

(b).

Clase verdadera	Asignado a la clase	
	SCO	SNC
SCO	33	0
SNC	0	33
% correcto de predicción	100	100
Habilidad de predicción 100%		

6.6.3. Regresión por mínimos cuadrados parciales (PLS)

Con el conjunto de espectros fue obtenida una regresión por mínimos cuadrados parciales (PLS); ecuación de predicción que asocie la variable HTP (mg kg^{-1} de suelo seco) como variable Y (matriz de referencia) y la variable X (matriz de datos espectrales); para ello se utilizó el software Pirouette Chemometrics Software versión 3.11 para Windows (Infometrix, 2003). Los espectros registrados del suelo Gleysol no contaminado y contaminado con petróleo crudo se muestran en la Figura 31.

El método de regresión de mínimos cuadrados (PLS) se sostiene sobre la relación del análisis de los componentes principales. Se emplea para encontrar relaciones fundamentales entre dos matrices (X e Y). Los modelos PLS tratan de encontrar direcciones multidimensionales en el espacio X que expliquen el máximo de varianza multidimensional en el espacio Y es especialmente práctico cuando la matriz de predicción tiene más variables que observaciones (Melgar, 2009).

PLS generaliza y combina características del análisis de componentes principales (PCA) y análisis de regresión múltiple. El PLS es utilizado para estudiar la estructura de covarianza entre los espacios correspondientes a dos matrices X y Y, para predecir un conjunto de variables dependientes a partir de un conjunto grande de variables independientes. Cada dirección en estos espacios es representada por una componente principal o variable latente, ya que la suposición básica de todos los modelos PLS, es que el sistema o proceso estudiado depende de un número pequeño de variables latentes (VL). Así, al finalizar los cálculos de PLS, se obtendrán cargas y puntuaciones para cada uno de los espacios X y Y (Nuñez *et al.*, 2009).

Para la obtención de la ecuación, la matriz de datos fue sometida previamente a una transformación matemática usando los tratamientos *Normalize*, *Smooth* y *Multiply*. La ecuación de calibración fue seleccionada sobre la base de minimizar el error estándar de calibración y de validación cruzada (SEC y SECV), y maximizar el coeficiente de correlación (RSQ) (Alomar y Fuchslocher, 1998). Estos parámetros proporcionan información a cerca de la bondad del modelo.

El objetivo de la calibración fue obtener parámetros que permitan calcular la concentración de HTP en futuras muestras, de forma que el residual (diferencia entre el valor de referencia y el predicho) de los datos calculados sea el menor posible. Teniendo en cuenta esto se evaluó la capacidad de predicción del modelo.

La elección del número de factores o variables latentes (componentes principales) se realizó mediante la validación cruzada, procedimiento en el cual el conjunto original se dividió en 9 grupos de cancelación, en la que se utiliza un grupo para comprobar los resultados (validar) y el resto para construir el modelo.

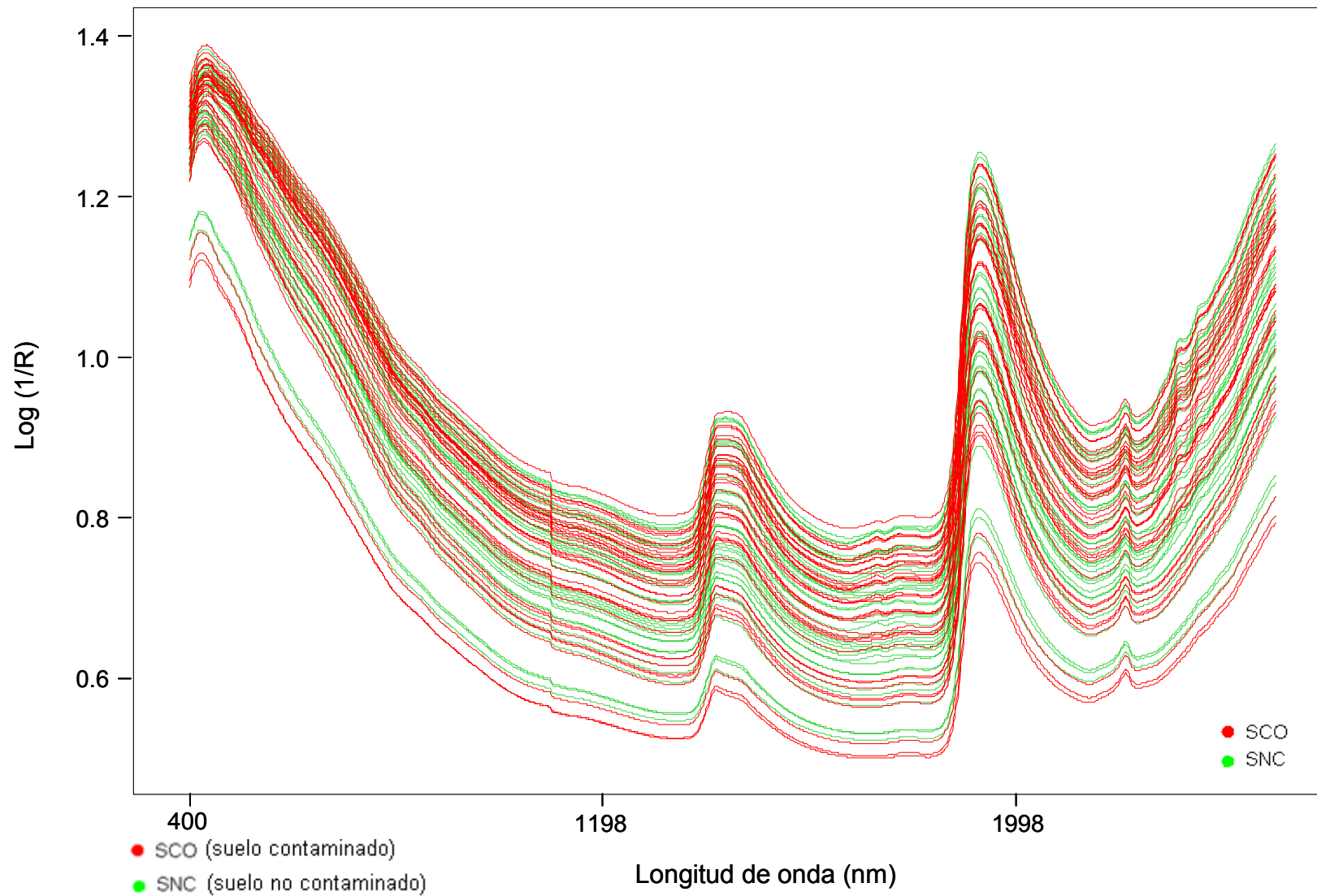


Figura 31. Espectro NIR de las muestras de suelo no contaminado y contaminado con petróleo crudo del suelo Gleysol.

6.6.4. Ecuaciones de Calibración

Los 1050 datos al ser graficados representan la huella espectral característica de cada muestra como se observa en la Figura 31. Al conjunto de 84 espectros obtenidos por NIR sobre suelo no contaminado y contaminado con petróleo crudo, se les realizó los tratamientos matemáticos que indicaron el número de componentes principales que presentaron la mayor información para la predicción de las muestras. Para la degradación de HTP en el tratamiento con suelo contaminado, la mayor variabilidad espectral explicada estuvo representada por el componente principal uno (PC1), que explicó el 93.79 % y el PC2 que explicó el 3.10 % de la variabilidad; en conjunto los PC1 y PC2 expresaron el 96.89 % de la variabilidad espectral.

Las ecuaciones obtenidas a partir del procedimiento estadístico PLS, *normalize*, *smooth*, *multiply* y nueve grupos de cancelación permitieron obtener estimaciones exactas en el proceso de degradación de petróleo crudo.

Las ecuaciones de calibración desarrolladas para todos los componentes por PLS, presentaron valores de los estadísticos de calibración satisfactorios, dado que el coeficiente de correlación (RSQ) fue superior a 0.9. Según Williams (2003), se definen siete niveles de precisión de la calibración basado en los valores RSQ. Los valores comprendidos entre 0.83 y 0.90 son útiles para la mayoría de las aplicaciones incluyendo la investigación, las predicciones aproximadas y el control de calidad.

En la Figura 32 es presentado el gráfico del modelo de calibración para el parámetro de HTP en suelo Gleysol no contaminado y contaminado con petróleo crudo, en este se observa que las 72 muestras presentaron un RSQ de 0.98, el cual de acuerdo con Williams (2003), es un valor aceptable.

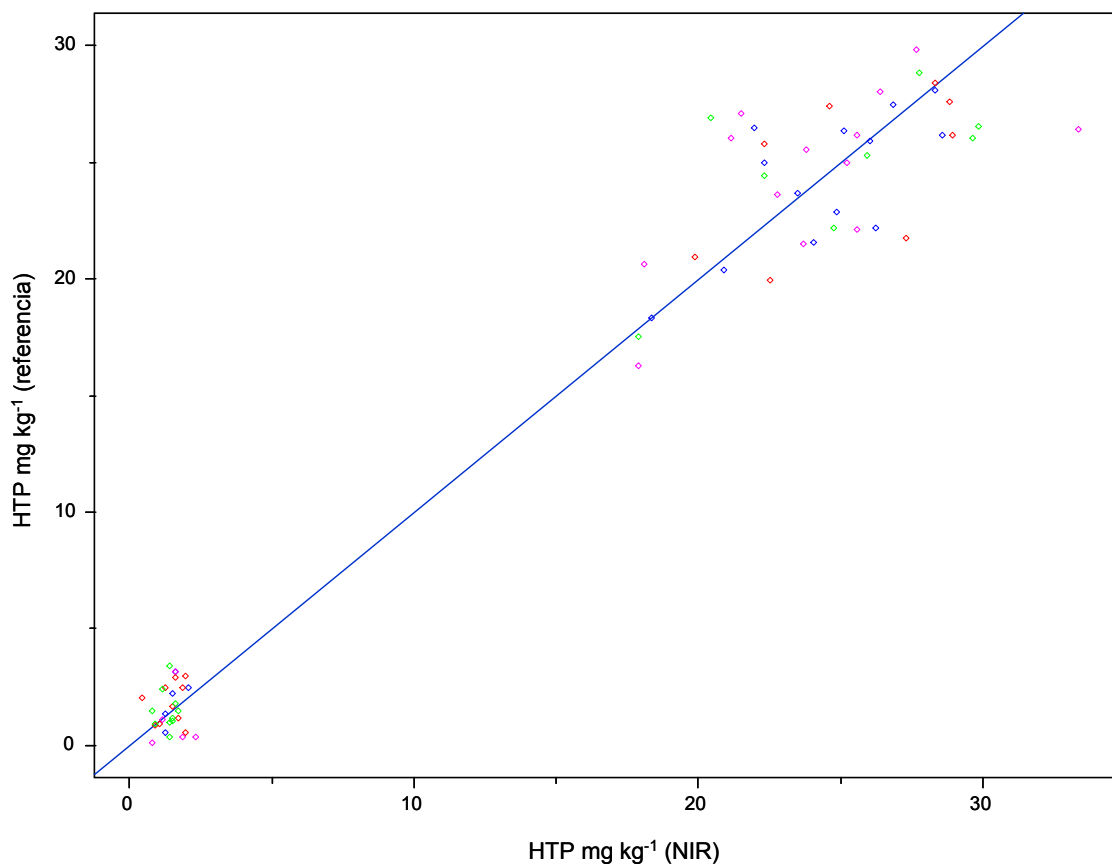


Figura 32. Modelo de calibración para HTP en suelo Gleysol no contaminado y contaminado con petróleo crudo.

6.6.5. Validación de la ecuación de calibración

El modelo de regresión (PLS) establecido para estimar la concentración de HTP en muestras de suelo, fue evaluada mediante validación interna y validación externa.

6.6.5.1. Validación interna

El modelo de estimación fue evaluado internamente por validación cruzada. El método consistió en dividir el conjunto de muestras en 9 grupos; 8 grupos se utilizaron para construir el modelo de calibración (el modelo se repitió tantas veces como grupos de cancelación hubieron) para el parámetro HTP mg kg⁻¹ de suelo seco y uno para comprobar resultados, como se muestra en el Cuadro 11. Donde N es el número de muestras usadas para obtener la ecuación de calibración, se muestran también los

tratamientos matemáticos para los parámetros evaluados. El modelo de calibración encontrado presenta un valor de RSQ de 0.997. En el cuadro son presentados los diferentes tratamientos matemáticos que se realizaron, los valores superior e inferior (rango) correspondiente a la muestras seleccionadas, los valores del coeficiente de correlación (RSQ) y el menor valor del error estándar de calibración y validación cruzada (SEC y SECV). La SECV se utiliza para indicar el número óptimo de factores de un modelo de regresión, siempre y cuando la curva presente una estructura clásica. Tomando en cuenta que cuando $SECV > SEC$, en las ecuaciones obtenidas, el modelo de calibración es válido para predecir muestras desconocidas (Infometrix, 2003).

En la Figura 33 es presentada la validación interna para el parámetro HTP estudiado en el suelo Gleysol, se observa que las muestras presentaron una buena relación entre las variables analizadas, teniendo un RSQ de 0.997; este valor indica un buen grado de asociación entre las variables, muestra el grado de explicación que el modelo estimado hace del problema que se analiza, en este caso la degradación de HTP leída en NIR en función a la lectura en Soxhlet. De acuerdo con Shenk y Westerhaus (1996), las calibraciones con valores de RSQ mayores a 0.9 poseen una excelente capacidad para predicciones cuantitativas, ya que existe una alta correlación entre los valores obtenidos en el laboratorio y los predichos mediante la técnica NIR.

Mediante el proceso de validación interna se obtienen los parámetros de calibración como SEC, BIAS y SEC(V). Al representar los valores de referencia frente a los predichos por NIR se obtienen los gráficos de validación interna para la concentración de HTP (Figura 33). Los resultados indican que es posible la cuantificación de los HTPs en el suelo Gleysol con un error de predicción (SEC) de 0.028.

En general, la ecuación propuesta es apropiada, no solo por el error estándar de calibración bajo, sino también por el alto valor del coeficiente de determinación (Vásquez *et al*, 2004). De acuerdo con estos resultados se confirma que es posible determinar la degradación de contaminantes, en este caso petróleo crudo en suelos contaminados mediante el uso de la tecnología NIRS.

Cuadro 11. Estimadores quimiométricos para la validación interna para HTP en suelo no contaminado y contaminado con petróleo crudo.

Parámetro	N	Tratamiento matemático	Rango	RSQ
HTP (mg kg ⁻¹ ss)	65	Normalize Smooth Multiply 2nd derivative	464-33333	0.997

SEC	SECV	No. de factores	Grupos de cancelación
962.53	825138.37	9	7

RSQ: Coeficiente de determinación; SEC: Error estándar de calibración; SECV: Error estándar de validación cruzada; N: Número de muestra.

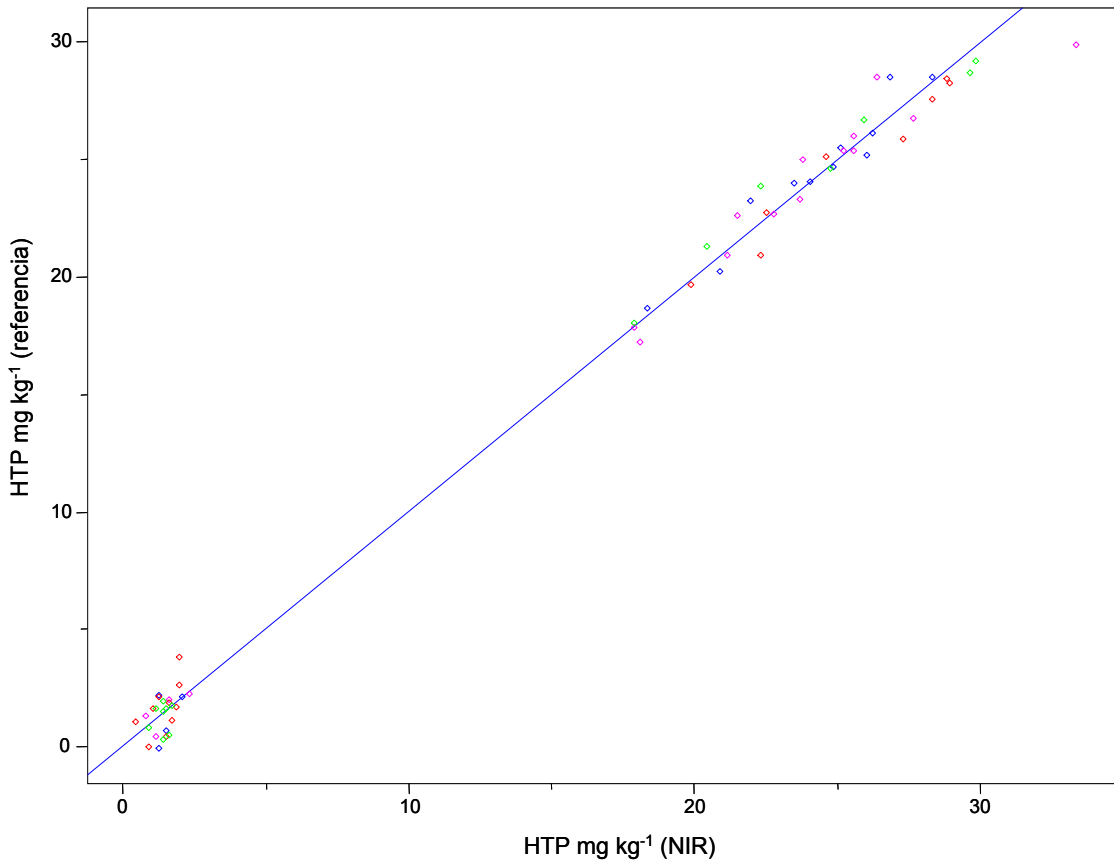


Figura 33. Comparación de los valores de referencia con los valores predichos por el modelo PLS en NIR para la determinación de HTP en suelo no contaminado y contaminado con petróleo crudo.

6.6.5.2. Validación externa

Para comprobar la robustez del modelo de calibración se realizó una validación externa con un conjunto de siete muestras de suelo que no pertenecen al colectivo de calibración; y que presentan un rango de concentración de HTP dentro de la aplicabilidad de las ecuaciones obtenidas en el Cuadro 12.

La capacidad del modelo fue evaluada usando la función de relación RPD, el cual es definido como la relación entre la desviación estándar del método químico de referencia y el error estándar (SD/SEP). Si esta relación es mayor que 2.5, entonces el modelo de calibración es adecuado y las ecuaciones resultan con elevado nivel de predicción (Murray, 1993).

Cuadro 12. Estimadores quimiométricos para la validación externa para HTP en suelo no contaminado y contaminado con petróleo crudo.

Parámetro	N	Tratamiento matemático	SD	Rango	RSQ
HTP (mg kg ⁻¹ ss)	7	Normalize Smooth Multiply 2nd derivative	13383.20	813-28571	0.98
		SEP	SEPC	No. de factores	RPD
		2585.21	13198.7	3	4.49

SD: Desviación estándar de datos de laboratorio; RSQ: Coeficiente de correlación; SEP: Error estándar de predicción; SEPC: Error estándar de predicción corregido; N: Número de muestra; RPD: Relación de desviación de la predicción (SD/SEP(C)).

El RSQ (0.98) obtenido es satisfactorio y se considera que posee una excelente capacidad para predicciones cuantitativas. De acuerdo con Campos *et al.* (2010), el RSQ ofrece información sobre la calidad de la calibración pero no sobre la precisión y exactitud de las predicciones. Según algunos autores (Park *et al.*, 1998; Murray, 1993)

la exactitud de las ecuaciones puede valorarse utilizando la relación SD/SEP donde SD es la desviación estándar de los valores de laboratorio de las muestras de calibración. Si este cociente es superior a 3 la calibración es robusta, si el valor se encuentra entre 2.5 y 3 la calibración es buena y si el cociente se encuentra entre 2.5 y 2, la calibración es aceptable para análisis cuantitativos y cualitativos respectivamente. Por su parte Williams y Sobering (1993), señalan que valores para la relación SD/SEP (RPD) superiores a 3, son adecuados para el empleo de la técnica NIR con criterios de diagnóstico e investigación. Este valor se considera ideal, si bien en los casos en los que SD_{ref} es menor que la unidad, puede no llegarse a ese valor ideal (Williams y Sobering, 1993).

De acuerdo con lo anterior y con resultado del RPD (4.49) obtenido en el análisis de suelo no contaminado y contaminado con petróleo crudo, se concluye que la técnica NIRS ofrece una alternativa viable para la determinación de petróleo crudo en suelos contaminados.

Los resultados proporcionados por la metodología NIR y los obtenidos por el método químico de referencia se compararon utilizando la prueba de t de Student para valores apareados (Ramis *et al.*, 2001). El nivel de confianza alcanzado aplicando dicha prueba fue de 1.20 para HTP en el suelo Gleysol no contaminado y contaminado (Cuadro 13). El nivel de confianza, para este parámetro, es mayor que 0.05 (elegido como el mínimo); lo que indica que no hay diferencias significativas entre los resultados obtenidos por ambos métodos. En este cuadro también se recogen otros dos parámetros como son el error cuadrático medio (*root mean standard error*) (RMSE) y la media de los residuales, indicadores del error que se puede cometer al predecir muestras desconocidas (Hernández, 2008). De acuerdo con estos resultados se tiene que no existe diferencia estadística significativa entre los métodos utilizados, al obtener un nivel de significancia (p) igual a 1.20, por lo que se concluye que ambos métodos son comparables. Este resultado se presenta en la Figura 34; en la cual, se correlacionan los valores obtenidos con el método de referencia y la predicción NIRS.

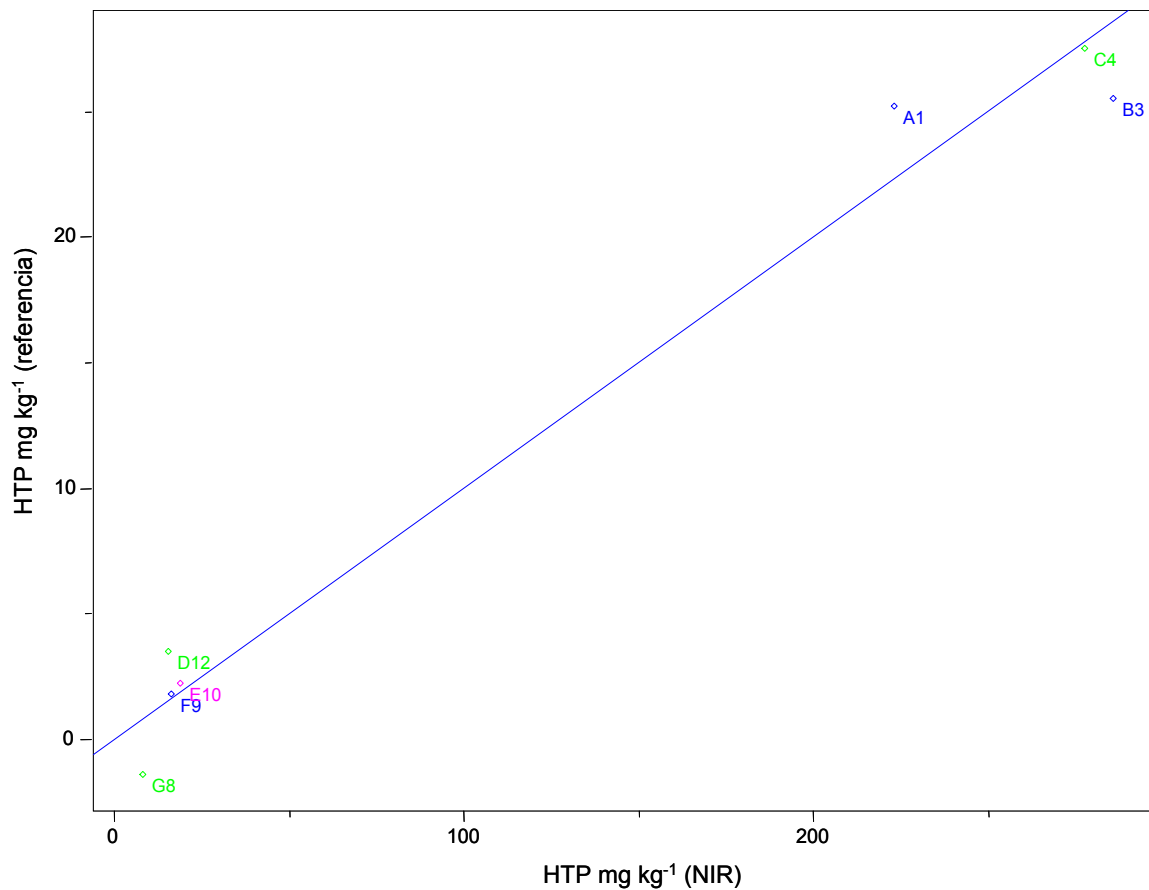


Figura 34. Validación externa del modelo de calibración para HTP en suelo Gleysol no contaminado y contaminado con petróleo crudo.

Cuadro 13. Validación externa del modelo de calibración mediante la prueba de t de Student.

Parámetro	p (nivel de significancia)	Media residual	RMSE
HTP	1.202	0.000096	2312.28

RMSE: Error cuadrático medio.

7. CONCLUSIONES

El presente trabajo el cual tuvo por objetivo evaluar la capacidad de la tecnología NIR para estudiar la degradación del petróleo crudo en suelo, y se concluye lo siguiente:

El uso de las técnicas quimiométricas no supervisadas (análisis exploratorio) como es la aplicación del análisis de componentes principales, en la interpretación de las señales generadas con la tecnología NIR, demostró ser adecuada para visualizar la distribución de las muestras en el espacio de los dos primeros componentes principales. Lo cual facilita determinar la técnica de clasificación y modelación, así como para proponer una solución rápida y eficaz en el estudio de la degradación de petróleo crudo en el suelo.

Con respecto al uso de las técnicas de reconocimiento de pautas supervisadas para clasificación y modelado de clases, el análisis discriminante lineal (LDA), resultó ser una técnica eficiente para la correcta clasificación y predicción de las muestras, es decir, es una técnica que permitió la generación de funciones discriminantes. En el presente trabajo el poder de clasificación fue del 100% y un poder de predicción también del 100%. Por lo que se concluye que el análisis discriminante lineal es una técnica útil para clasificar y predecir muestras de suelo contaminado con petróleo crudo del que no lo es, a partir de la información contenida en las huellas espectrales.

En cuanto a la técnica de modelado suave independiente de analogía clases (SIMCA), se concluye que es una herramienta útil para construir modelos de clasificación y predicción a partir de las huellas espectrales NIR; en los resultados de este estudio se obtuvo un poder de clasificación del 100% y un poder de predicción del 100%. Por lo que resulta útil para clasificar y predecir una muestra de suelo contaminada por petróleo crudo.

Los modelos PLS desarrollados con los espectros NIR obtenidos durante el proceso de degradación de petróleo crudo en suelos; cumplen con los indicadores quimiométricos de la evaluación interna y externa. Con lo cual, se tiene una alternativa para cuantificar

el nivel de contaminación con TPH en suelos, con resultados similares a los obtenidos con el método convencional de referencia.

La espectroscopia en el infrarrojo cercano es una herramienta que genera información suficiente y útil que permitió realizar modelos de clasificación y predicción de la degradación de petróleo crudo en un suelo contaminado, con resultados similares a los obtenidos con el método de referencia.

8. LITERATURA CITADA

Acuña, O., Peña, W., Serrano, E., Pocasangre, L., Rosales, F., Delgado, E., Trejos, J., y Segura, A. 2006. La importancia de los microorganismos en la calidad y salud de suelos. XVII Reunião Internacional da Associação para a Cooperação nas Pesquisas sobre Banana no Caribe e na America Tropical, Santa Catarina, Brasil. p. 222-233.

Adams, S. R. H., Domínguez, R. V. y García H. L. 1999. Potencial de la biorremediación de suelo y agua impactados por petróleo en el trópico mexicano. *Terra*, 17:159-174.

Agilent Technologies. 2000. Cromatógrafo de gases serie agilent 6890. Wilmington Delaware, USA. Analytical Methods Committee. *Analyst* 1994, 119, 2363.

Alexander, M. 1990. Introducción a la microbiología del suelo. AGT Editor, México. 491p.

Alexander, M. 1999. Biodegradation and bioremediation. Academic Press. USA. 453p.

Alomar, D., y Fuchslocher, R. 1998. Fundamentos de la espectroscopía de reflectancia en el infrarrojo cercano (NIRS) como método de análisis de forrajes. *Agrosur*, 26(1):88-104.

Alomar, D., y Fuschlocher, R. 1998. Fundamentos de la espectroscopía de reflectancia en el infrarrojo cercano (NIRS) como método de análisis de forrajes. *Agrosur*, 26(1):88-104.

Atlas, M. R., y Bartha, R. 2002. Ecología microbiana y microbiología ambiental. 4ª edición. Editorial Addison Wesley, España. p.675.

Atlas, R. M. 1977. Stimulated petroleum biodegradation. *CRC Crit. Rev. Microbiol*, 5:371-386.

Atlas, R. M. 1978. Microorganisms and petroleum pollutants. *BioScience*, 28:387-391.

Atlas, R. M. 1981. Microbial degradation of petroleum hydrocarbons: An environmental perspective. *Microbiological Reviews*, 45(1):180-209.

Barnes, R.J., Dhanoa, M. S., and Lister, S. J. 1989. Standard normal variate transformation and detrending of near infrared diffuse reflectance spectra. *Appl. Spectrosc.* 43:772-777.

Batten, G. D. 1998. Plant analysis using near infrared reflectance spectroscopy: the potential and the limitations. *Australian Journal of Experimental Agriculture*, 38: 697-706.

Bautista, M. M. 2009. Avances en la aplicación de la espectroscopia NIR en la industria farmacéutica. Introducción a PAT y técnicas de imagen. Tesis doctoral en Química.

Universidad Autónoma de Barcelona, España. 203p.

Bedair, H. M., and Al-Saad, H. T. 1992. Dissolved and particulate adsorbed hydrocarbons in water of shaft al Arab river, Iraq. *Water, Air, Soil Pollution* 61: 397-408.

Beebe, K. R., Pell, R. J., and Seasholtz, B. M. 1998. Chemometrics: A practical guide. New York, Chichester, Weinheim, Brisbane, Singapore, Toronto.

Ben-Dor, E. And Banin, A. 1995. Near infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Sci. Soc. Am. J.*, 59:364-372.

Blanco, M., and Villarroya, I. 2002. NIR spectroscopy: a rapid-response analytical tool. *Trends in Analytical Chemistry*, 21:240-250.

Blanco, M., y Cerdá, V. 1988. Quimiometría. Publicaciones de la Universidad Autónoma de Barcelona. Bellaterra.

Boldu, F., Vervoort, J., Grontehuis, J. y Van Groenestijn, J. 2002. Substrate interactions during the biodegradation of Benzene, Toluene Ethylbenzene and Xylene (BTEX) Hydrocarbons by the fungus *Cladophialophora* sp. strains T1. *Applied and Environmental Microbiology*, 68:2660-2665.

Botello, V. A. 1996. Características, composición y propiedades fisicoquímicas del petróleo. In A. V. Botello, L. J. Rojas G., A. Benítez J. y L. D. Zárata (eds.). Golfo de México. Contaminación e impacto ambiental: Diagnostico y Tendencias. Programa EPOMEX. Universidad Autónoma de Campeche, Méx. pp. 204-210.

Bouveresse, E., Massart, D. L. 1996. Standardisation of Near Infrared Spectrometric Instruments: a *Review Vib. Spectroscopy*. 11-3.

Brereton, G. R. 2002. Chemometric, data analysis for the laboratory and chemical plant. John Wiley & Sons, Ltd. Chichester. England.

Brereton, R. G. 1990. Chemometrics applications of mathematics and statistics to laboratory systems. Ellis Harwood, Chichester. England.

Brereton, R. G. 2000. Introduction to multivariate calibration in analytical chemistry. *Analyt tutorial review*, 125:2125-2154.

Brereton, R. G. 2003. Chemometrics data analysis for the laboratory and chemical plant. John Wiley & Sons, Ltd. Chichester. England.

Brzezińska, M., Stepniewska, Z., and Stepniewski, W. 1998. Soil oxygen status and dehydrogenase activity. *Soil Biol. Biochem.* 30(13):1783-1790.

Camacho, C. O., del Valle, P. D., y Ruelas, A. G. A. 1992. SAS (Statistical Análisis

System) para microcomputadoras. Colegio de Postgraduados. Pp. 174.

Camel, V. 2002. Extraction techniques. *Analytical and Bioanalytical Chemistry*, 2:39-40.

Campo, L., Castro, P., y Moreno-González, J. 2010. Ecuaciones de calibración preliminares para la evaluación de la calidad de la biomasa en plantas de maíz por NIRS. Congreso SEEP y 4ª Reunión Ibérica de Pastos y Forrajes, 3-6 mayo 2010. pp.135-139.

Carvalho, A. R., M. Noga Sanchez, J. Wattoon, and R. G. Brereton. 2005. Comparison of PLS and kinetic models for a second-order reaction as monitored using ultraviolet visible and mid-infrared spectrometry. *Talanta*, 68(4):1190-200.

Cerisola, C. I., García, M. G., y Filgueira, R. R. 2005. Distribución de la porosidad de un suelo franco arcilloso (Alfisol) en condiciones semiáridas después de 15 años bajo siembra directa. *Cienc. Suelo*, 23(2):167-178.

Cerniglia, C. E. 1984. Microbial metabolismo of polycyclic aromatic hydrocarbons. *Advances in Applied Microbiology*, 30:31-71.

Chakraborty, S., Weindorf, C. D., Morgan, L. S. C., Ge, Y., John M. Galbraith, M. J., Li, B., and Kahlon, S. Ch. 2010. Rapid identification of oil-contaminated soils using visible near-infrared diffuse reflectance spectroscopy. *J. Environ. Qual.* 39:1378-1387.

Chodak, M., Niklińska, M., and Beese, F. 2007. Near-infrared spectroscopy for analysis of chemical and microbiological properties of forest soil organic horizons in a heavy-metal-polluted area. *Biology and Fertility of Soils*, 44(1):171-180.

CIMADES (Comisión Interinstitucional del Medio Ambiente y Desarrollo Social. 1999. Contingencias ambientales en el estado de Tabasco, 1993-1998. Dirección de Protección Ambiental CIMADES. Villahermosa, Tabasco, México.

Coimbra, M. A., Gonscalves, F., Barros, A. S. y Delgadillo, I. 2002. Fourier transforms infrared spectroscopy and chemometric analysis of white wine polysaccharide extract. *J. Agric. Food Chem.* 50:3405-3411.

Coomans, D., Massart, D. L., and Kaufman, L. 1979. Optimization by statistical linear discriminate analysis in analytical chemistry. *Analytica Chimica Acta*, 112:97-122.

Coyne, M. 2000. Microbiología del suelo: un enfoque exploratorio. Paraninfo, España. 416p.

Cozzolino D. 2002. Uso de la espectroscopia de reflectancia en el infrarrojo cercano (NIRS) en el análisis de alimento para animales. *Agrociencia*, 6:123-132.

Cozzolino, D., Restaino, E., La Manna, A., Fernandez, E., and Fassio, A. 2009.

Usefulness of near infrared reflectance (NIR) spectroscopy and chemometrics to discriminate between fishmeal, meat meal and soya meal samples. *Cien. Inv. Agr.* 36(2):209-214.

Cuadrenys, O. J. 1981. Formulación y nomenclatura II Química Orgánica. Edit. Teide, Barcelona, España.

Davies, A. M. C., and Williams, Ph. 1996. Near infrared spectroscopy: The future waves. NIR Publications, Chichester, Reino Unido.

Dea, J. R. 1998. Extraction methods for environmental analysis. Johns Wiley & Sons. Chichester, England.

Dhanao, M. S., Lister, S. J., Sanderson, R., and Barnes, R. J. 1994. The link between Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) transformations of NIR spectra. *J. Near Infrared Spectrosc.* 2:43-47.

Diario oficial de la federación (DOF). 2005. Norma oficial mexicana NOM-138-SEMARNAT-2003. Límites máximos permisibles de hidrocarburos en suelos y las especificaciones para su caracterización y remediación.

Doménech, X. 1999. Química de la contaminación. Miraguano Ediciones. Madrid, España. 158p.

Doran, J. W., and Parkin, T. 1994. Defining and assessing soil quality. *In: Doran, J. W., Coleman, D. C., Bezdicek, D. F., Stewart, B. A. (eds.). Defining soil quality for a sustainable environment. Soil Sci. Soc. Am., Am. Soc. Agron., Madison, WI. p 3-21.*

Dorn, B. P., Vipond, E. T., Salanitro, P. J., and Wisniewski, L. H. 1998. Assessment of the acute toxicity of crude oils in soils using earthworms, microtox, and plants. *Chemosphere* 37: 845-860.

Downey, G. 1998. Food and Food Ingredient authentication by Mid- Infrared Spectroscopy and Chemometrics. *Trends Anal. Chem*, Vol. 17.

Edelmann, A., Diewok, J., Rodríguez, J. and Lendl, B. 2003. High-performance liquid chromatography with diamond ATR-FTIR detection for the determination of carbohydrates, alcohols and organic acids in red wine. *Analytical and Bionalytical Chemistry.* 376:92-97.

Ensley, B. D., y Zylstra, G. J. 1997. principios y prácticas de biotratamiento utilizando microorganismos modificados. *In: Biotratamiento de residuos tóxicos y peligrosos.* 1997. Levin, M., y Gealt, M. A. (eds.). McGrawHill. 338p.

Ercoli, E., Saracino. C., Galvez, J., y Videla. O. 1998. Evaluación de dos años de experiencias de campo en biorremediación de suelos. 3ras Jornadas de Preservación

de Agua, Aire y Suelo en la Industria del Petróleo y del Gas, IAPG: Tomo 1, 47-60.

Erikson, M., Fäldt, J., Dalhammar, G., and Borg-Karlson, A. K. 2001. Determination of hydrocarbons in old creosote contaminated soil using headspace solid phase microextraction and GC-MS. *Chemosphere*, 44:1641-1648.

Etchevers, J. D. 1988. Análisis químico de suelos y plantas. Centro de Edafología. Colegio de Posgraduados, Chapingo, Estado de México. 803p.

Etchevers, J. D. 1992. Análisis de suelos. *In: los análisis físicos y químicos*. 1992. Alcántar, G. G., Etchevers, B. J. D., Aguilar, S. A. (eds.). Colegio de Postgraduados, Centro de Edafología. 125p.

Eweis, J. B., Ergas, D. J., Chang, D. P. Y., y Schroeder, E. D. 1999. Principios de biorrecuperación. Tratamientos para la descontaminación y regeneración de suelos y aguas subterráneas mediante procesos biológicos y fisicoquímicos. McGrawHill, México, pp. 327.

FAO. 1990. FAO/UNESCO: Mapa mundial de suelos. Leyenda revisada. Informe sobre recursos mundiales de suelos, No. 60. FAO. Roma, Italia. 125p.

FAO. 2007. Base referencial mundial del recurso suelo. Marco conceptual para la clasificación, correlación y comunicación internacional. Reporte No. 103. IUSS, ISRIC, FAO. Roma, Italia. 87p.

Fenchel, T., King, G. M., and Blackburn, T. H. 2000. Bacterial Biogeochemistry. The Ecophysiology of Mineral Cycling. 2°ed. Academic Press, San Diego. pp. 43-59, 117-161.

Fernández, C. P., Martín-Esteban A., Pérez-Conde, C., y Vidal, M. 2002. Toma y tratamiento de muestras. Editorial Síntesis S. A. Madrid, España.

Fernández, L., Rojas, N., Roldan, T., Ramírez, M., Zegarra, R., y Reyes, R. 2006. Manual de técnicas de análisis de suelos aplicadas a la remediación de sitios contaminados. SEMARNAT-INE-IMP. México.

Fernández, P., Carrera, G., Grimalt, J. O., Ventura, M., Camarero, L., Catalán, J., Nickus, U., Thies, H., and Psenner, R. 2003. Factors governing the atmospheric deposition of polycyclic aromatic hydrocarbons to remote areas. *Environmental Science & Technology*, 37: 3261-3267.

Fernández-Linares, L. 2000. Tecnologías de biorremediación en suelos. p. 700-705. In: Quintero-Lizaola R. Reyna-Trujillo T, Corlay-Chee L., Ibáñez-Huerta A. y García-Calderón (eds.). La edafología y sus perspectivas al siglo XXI. Tomo II Colegio de Postgraduados, Universidad Autónoma de México, Universidad Autónoma Chapingo, México 700-705 p.

Flores, P. M. A., Torras, O. S., y Téllez, G. R. 2004. Medidas de mitigación para uso de suelos contaminados por derrames de hidrocarburos en infraestructura de transporte terrestre. Publicación Técnica No 257 Sanfandila, Qro., México.

Font, V. L., Calero, M. B., del Castillo, R. A., y Valenciano, M. 2003. Efecto del nitrógeno sobre el estado microbiológico de un suelo ferrítico rojo en un agroecosistema citrícola, *Centro Agrícola*, 30(1):5-9.

Gallardo-Velázquez, T., Osorio-Revila, G., Zúñiga-de Loa M. y Rivera-Espinoza. 2009. Application of FTIR-HATR spectroscopy and multivariable analysis to the quantification of adulterants in Mexican honeys. *Food Res. Int.* 42(3):313-318.

García, Ch. M. E. 2009. Diseño e implementación de una herramienta de detección facial. Tesis de maestría en tecnología de cómputo. Instituto Politécnico Nacional. México D.F. 99p.

García, S. A., Ramos, M. N., y Ballesteros, E. 2005. Estudio comparativo de distintas técnicas analíticas (espectroscopía de NIR y RMN y extracción mediante Soxhlet) para la determinación del contenido graso y de humedad en aceitunas y orujo de Jaén. *Grasas y Aceites*, (220)56-3:220-227.

García-López, E., Zavala-Cruz, J. y Palma-López, D. J. 2006. Caracterización de las comunidades vegetales en un área afectada por derrames de hidrocarburos. *TERRA Latinoamericana*, 24(1):17-26.

García-Rivero, M., y Peralta-Pérez, M. R. 2008. Cometabolismo en la biodegradación de hidrocarburos. *Revista Mexicana de Ingeniería Química*, 7(1):1-12

Garrett, R. M., Guenette, Ch. C., Haith, K. E., and Prince, R. C. 2000. Pyrogenic polycyclic aromatic hydrocarbons in oil burn residues. *Environmental Science & Technology*. 34: 1934-1937.

Gee, G. W., and Bauder, J. W. 1986. Particle-size analysis. Methods of soil analysis. Part 1. Physical and mineralogical methods. American Society of Agronomy, Inc., and Soil Science Society of America, Inc., Madison, Wis. pp. 383-411.

Gibson, D. T. 1982. Microbial degradation of hydrocarbons. *Environ. Toxicol. Chem.* 5:237-250.

Gibson, D. T., and Perales, R. E. 2000. Aromatic Hydrocarbon dioxygenases in environmental biotechnology. *Curr. Op. Biotech.* 11; 236-246.

González, M. I., González, P. C., Hernández, M. J., and Álvarez, G. N. 2003. Determination of fatty acids in the subcutaneous fase of Iberian breed swine by near infrared spectroscopy (NIRS) with a fibre-optic probe. *Meat Science*, 65:713-719.

Goodacre, R., and Kell, D. B. 1996. Correction of mass spectral drift using artificial neural networks. *Analytical Chemistry* 68:271-280.

Goyal, S. M., and Gerba, C. P. 1979. Comparative adsorption of human enteroviruses, simian rotovirus and selected bacteriophages to soils. *Applied Environment Microbiology* 38, 241-247.

Gros, A., y Domínguez, A. 1992. Abonos guía práctica de la fertilización. 8va. edición. Ediciones Mundi-Prensa. Madrid. 450p.

Guerrero, P. A. 2006. Estudio de la contaminación en suelos por petróleo crudo mediante generación de espacio de Cabeza-Espectrometría de masas. Tesis de Doctorado. Universidad de Salamanca. Salamanca, España.

Heichen, J. R., Bottomley, P. P., Cromack, J., and Myrold, D. 2000. Community composition and functioning of denitrifying bacteria from adjacent meadow and forest soil. *Applied and Environmental Microbiology*, 69(10):5974-5982.

Hernández, H. J. M. 2008. Utilización de la tecnología NIRS con sonda de fibra óptica para el control de calidad de Alfalfa. Tesis de Doctor en Química. Universidad de Salamanca, España. 116p.

Hopke, K. P. 2003. The Evolution of Chemometrics *Review Anal. Chem. Acta.* 500:365-377.

INEGI. Instituto Nacional de Estadística Geográfica e Informática. 2000. Anuario Estadístico del Estado de Tabasco. México D.F.

Infante, G. S., y Zárate de Lara, G. P. 2007. Métodos estadísticos. Un enfoque interdisciplinario. Trillas, México. 643p.

Infometrix, 2003. Pirouette. Multivariate data analysis. Ver. 3.11. Woodinville, W. A. USA.

Instituto Mexicano del Petróleo (IMP). 2008. Acerca del petróleo. Tipos de petróleo. <http://imp.mx/petroleo/apuntes/tipos.htm>. Consulta: 23 de agosto de 2010

Instituto Mexicano del Petróleo (IMP). 2008. El petróleo. <http://www.imp.mx/petroleo>. Consulta: 18 de agosto del 2010.

Irundayaraj, J., Sivakesava, S., Kamath, S. and Yang, H. 2001. Monitoring chemical changes in some foods using Fourier Transform Photoacoustic Spectroscopy. *Journal of Food Science.* (66)9:1416-1421.

Janik, L.J., Merry, R.H., and Skjemstad, J. O. 1998. Can mid infrared diffuse reflectance analysis replace soil extractions. *Australian Journal of Experimental Agriculture.* 38:681-

696.

Jarquín, S. A. 2006. Determinación de materia orgánica y nitrógeno en suelos tropicales por espectroscopía de infrarrojo cercano y quimiometría. Tesis de maestría, Colegio de Postgraduados, Campus Tabasco, México, 129p.

Jenkinson, D. 1992. La Materia Orgánica del suelo: Evolución. En: Wild A. Condiciones del suelo y desarrollo de las plantas. Madrid. Mundi-Prensa.

Julca, O. A., Meneses, F. L., Blas, S. R., y Bello, A. S. 2006. La materia orgánica, importancia y experiencias de su uso en la agricultura, IDESIA, 24(1):49-61.

Kesley, J., Kottler, B., and Alexander, M. 1997. Selective chemical extractants to predict bioavailability of soli-aged organic chemicals. *Environ. Sci. Technol.* 31: 214-217.

LaGrega, M. D., Buckingham, P. L., and Evans, J. C. 1996. Gestión de residuos tóxicos. tratamiento, eliminación y recuperación de suelos. Vol. I y II. McGraw-Hill. Madrid, España. Pp. 1316.

Lambert, P., Fingas, M., and Goldthorp, M. 2001. An evaluation of field total petroleum hydrocarbon (TPH) systems. *Journal of Hazard Materials*, 83:65-81.

Leahy, G. J., and Colwell, R. R. 1990. Microbial Degradation of Hydrocarbons in the Environment. *American Society for Microbiology*, 54(3):305-315.

Levin, A. M., y Gealt, A. M. 1997. Biotratamiento de residuos tóxicos y peligrosos. McGraw-Hill, México. 338pp.

Luna, G. M. L., Vega, J. C., Franco, H. M. O., Vásquez, M. S., Trujillo, T. N., Ramírez, F. E., y Dendooven, L. 2002. Actividad microbiana en suelos. *Avance y Perspectiva*, 21:328-332.

Luque-García, J. L., and Luque, de C. M. D. 2001. Water Soxhlet extraction assisted by focused microwaves: a clean approach. *Analytical Chemistry*, 73:5903-5908.

Macho, S. and M. S. Larrechi. 2002. Near-infrared spectroscopy and multivariate calibration for the quantitative determination of certain properties in the petrochemical industry. *Trends in Analytical Chemistry*, 21:799-806.

Macías-Rodríguez, L. E., Quero-Gutiérrez, A. y López, M. G. 2004. Caracterización de tres cultivares de fresa (*Fragaria xananassa* Duch) por espectroscopia de infrarrojo medio y quimiometría. *Agrociencia*, 38:487-495.

Malley, F. D., Hunter, N. K., and Webster, B. R. G. 1999. Analysis of diesel fuel contamination in soils by near-infrared reflectance spectrometry and solid phase microextraction-gas chromatography. *Soil and Sediment Contamination: An*

International Journal, 8(4):481-489.

Manahan, E. S. 2007. Introducción a la química ambiental. Reverté, México. 603 pp.

Martens, H., Næs, T. 1996. Multivariate calibration, Jhon Wiley & Sons, Chichester.

Martínez, E. V., y López, S. F. 2001. Efecto de hidrocarburos en las propiedades físicas y químicas de suelo arcillosos. *TERRA Latinoamericana*, 19(1):9:17.

Massart, D. L., Vandeginste, B. G. M., Deming, S. M., Michote, Y., and Kaufman, L. 2003. Chemometric: a textbook. ELSEVIER, Amsterdam. 488p.

Massart, D. L., Vandeginste, B. G. M., Buydens, L. M. C., De Jong, S., Lewi, P. J., and Smeyers-Verbeke, J. 1997. Handbook of Chemometrics and Qualimetrics. Elsevier. Amsterdam. 876 p.

Melgar, L. M. G. 2009. Desarrollo de un método quimiométrico acoplado FTIRHATR para la determinación de las principales propiedades químicas del queso panela. Tesis de Maestro en ciencias de alimentos. Instituto Politécnico Nacional, México. 157p.

Miller, C. E. 2005. Chemometrics in Process Analytical Chemistry. *In*: Bakeev, K. (ed.) Process Analytical Technology. Blackwell publishing, Oxford, pp 226-227.

Miller, J. N., and Miller, J. C. 2000. Statistics and chemometrics for analytical chemistry. Pearson Education, Harlow.

Montgomery, C. D., Peck, E. A., y Geoffrey, V. G. 2005. Introducción al análisis de regresión lineal. 3ª edición. CECOSA. México. 588p.

Montoliu, I. 2001. Aplicación y desarrollo de técnicas quimiométricas para su aplicación en la Industria de Alimentos. Tesis doctoral. Universitat Autònoma de Barcelona, España.

Morer, X. T. 2003. La quimiometría en el análisis de datos medioambientales. IX Conferencia Española de biometría. p23-30.

Murray, I. 1993. Forage analysis by near infrared spectroscopy. Pages 285–312 *in* Davies, A., Baker, R. D., Grant, S. A., and Laidlaw, A. S. (eds.). Sward herbage measurement handbook. British Grassland Society, University of Reading, Reading, UK.

Mustin, M. 1987. Le compost. Ed. François Dubusc. París, 954 p.

Navarro, P. J., Herrero, G. M. L., y Beneyto, M. 1995. Residuos orgánicos y agricultura. Universidad de Alicante. Servicio de Publicaciones. Alicante. España, 108 p.

Ney, R. 1990. Fate and transport of organic chemicals in the environment. Govern. Inst.

Maryland. USA.

Nikolich, K., Sergides, C., and Pittas, A. 2001. The application of Near Infrared Reflectance Spectroscopy (NIRS) for the quantitative analysis of hydrocortisone in primary materials. *J. Serb. Chem. Soc.* 66(3):189-198.

Norris, K. H., and Williams, P. C. 1984. Optimization of mathematical treatments of raw near-infrared signal in the measurement of protein in hard red spring wheat: I. Influence of particle size. *Cereal Chem.* 61:158-165.

Núñez, O., Porro, D., Talavera, I., Bustio, L., Hernández, N., y Larín, R. 2009. Nuevo sistema automatizado para el análisis de datos químicos y bioquímicos. VII Congreso Internacional de Informática en la Salud, La Habana, Cuba.

Ortiz, C. A. 1999. Biodegradación de petróleo crudo (Maya) en una matriz sólida, en reactores tipo columna. Tesis de Licenciatura. Cuernavaca, Morelos. 73p.

Osborne, B. G., Fearn, T. and Hindle, P. H. 1998. Practical NIR spectroscopy with applications in food and beverage analysis. Longman Scientific & Technical. 2nd ed. UK. 227p.

Palma-López D. J., Cisneros D. J., Moreno C.E., y Rincón-Ramírez J. A: 2007. Suelos de Tabasco: Su uso y manejo sustentable. Colegio de Postgraduados-ISPROTAB-FUPROTAB. Villahermosa, Tabasco, México.

Palma-López, D. J., Obrador, J. J., Zavala, C. J., García, L. E., Ortiz, G. C. F., Camacho, Ch. W., Meseguer, E. R., Sol, S. A., Juárez, L. F., Jasso, M. J., Guerrero, P. A., Ramos, R. R., y Triano, S. A. 1999. Diagnóstico de los recursos naturales, niveles de contaminación y alternativas para el desarrollo del área de influencia de los campos petroleros Cinco Presidentes y la Venta Norte. Campus Tabasco, Colegio de Postgraduados, Gobierno del Estado de Tabasco. H. Cárdenas, Tabasco. 353p.

Park, R. S., Agnew, R. E., Gordon, F. J., and Steen, R. W. J. 1998. The use of near infrared reflectance spectroscopy (NIRS) on undried samples of grass silage to predict chemical composition and digestibility parameters. *Animal Feed Science and Technology*, 72:155-167.

Pelletier, E., Ouellet, S., and Paquet, M. 1991. Long-term chemical and cytological assessment of oil contamination in estuarine intertidal sediments. *Marine Pollution Bulletin*, 22:273-281.

Pérès, C., Viallon C., and Berdagué J. L. 2002. Curie point pyrolysis-mass spectrometry applied to rapid characterization of cheeses. *Journal of Analytical and Applied Pyrolysis*, 65:161-171.

Perrotta, G. and Xefferis, P. 1996. Small synthetic aperture radar satellite constellations

for tracking oil spills. *Spill Science & Technology Bulletin*, 3:73-82.

Physical/chemical methods. 1986. Test Methods for Evaluating Solid Waste. USEPA publication (SW-846 on-line). Third edition. Environmental Protection Agency. United State of America.

PROFEPA, 2000. Disposición y procedimientos para la caracterización y restauración de suelos contaminados. Memoria GDT (1998-2000) sobre restauración de suelos contaminados. México, D. F. 147 p.

PROFEPA. 2010. Reporte de Emergencias ambientales en México. Procuraduría Federal de Protección al Ambiente. México. <http://www.profepa.gob.mx/PROFEPA/EmergenciasAmbientales/AnalisisNacionaldeEmergenciasAmbientales/>. Consulta: 19 de agosto del 2010.

Radovic, B. S., Goodacre R., and Anklam, E. 2001. Contributions of pyrolysis-mass spectrometry (Py-MS) to authenticity testing of honey. *Journal of Analytical Applied Pyrolysis*, 60:79-87.

Raina, M. M., Pepper, I. L., and Gerba, Ch. P. 2000. Environmental microbiology. Academic Press. 585 p.

Ramis, R G., y García-Álvarez, C. M. C. 2001. Quimiometría. Síntesis, Madrid.

Razo, E. D. I. 2003. Biorremediación de suelos contaminados con fenol. Tesis de Licenciatura en Quimicofarmacobiología. Universidad de las Américas, Puebla, México.

Rebolledo, R. H. H. 2002. Manual SAS por computadora. Análisis estadístico de datos experimentales. Trillas, México. 208p.

Reyes, S. C. A. 2009. Tipificación de miel de diferente origen botánico o geográfico mediante la espectroscopia del infrarrojo cercano. Tesis de maestría, Colegio de Postgraduados, México. 96p.

Reyes, C. P. 1985. Bioestadística aplicada. Trillas, México. 216p.

Reyes, C. P. 2003. Diseños experimentales. Trillas, México. 348p.

Riser-Roberts, E. 1992. Bioremediation of petroleum contaminated sites. C. K. Smoley, Boca Raton, FL, USA. 197p.

Rivera, C. M. del C. 2001. Microorganismos rizosféricos de los pastos alemán [*Echinochloa polystachya* (H.B.K.) Hitchc.] y cabezón (*Paspalum virgatum* L.) en la degradación del petróleo crudo y el benzo(a)pireno. Tesis de Doctorado en Ciencias en Edafología. Colegio de Postgraduados. Montecillo, edo. México. 348p.

Rivera-Cruz, M. del C., Ferrera-Cerrato, R., Volke-Haller, V., Fernández-Linares, L y Rodríguez-Vázquez, R. 2002. Poblaciones microbianas en perfiles de suelos afectados por hidrocarburos del petróleo en el estado de Tabasco, México. *Agrociencia*, 36:149-160.

Rodríguez, F. H., y Rodríguez, A. J. 2002. Métodos de análisis de suelos y plantas. Criterios de interpretación. Trillas. México. 196 pp.

Rosenberg, E and Ron E. Z. 1996. Bioremediation of petroleum contamination. *In*: Crawford, R. L., and Crawford, D. L. (eds.). Bioremediation principles and applications. Biotechnology Research Series 6 J. Lynch (ed) Cambridge. University Press. Cambridge, UK. pp 100-124.

Sabiroba, I. 2000. The potential application of halotolerant microorganisms for bioremediation of soils polluted with crude oil and hypersaline wastewaters. Kazan State University.

SAS Institute Inc. 2002. SAS/IML* Software: Usage and Reference, version 9, SAS Institute Inc. Cary, N C.

Savitzky, A., and Golay, M. J. E. 1964. Smoothing and differentiation of data by simplified least-squares procedures. *Anal. Chem.* 36(8):1627-1639.

Schmidt, D. E. 1999. Historical overview of oil spills from all sources (1960-1998). International Oil Spill Conference. 8-11 March. Seattle, USA.

Schmidt, D. E. 2001. Analysis of oil spill trends in the United States and worldwide. International Oil Spill Conference. 26-29 March. Tampa, Florida, USA.

Sheller, U., Zimmer, T., Becher, D., Schauer, F. and Shunck, W. H. 1998 Oxygenation cascade in conversion of *n*-alkanes to α,ω -Dioic acids catalyzed by Cytochrome P450 52A3. *J. Biol. Chem.* 273:528-534.

Shenk, J. and Westerhaus, M. O. 1996. Calibration the ISI way. *In*: "Near Infrared Spectroscopy: The future Waves". Ed. A.M.C. Davies and Phil Williams. 198-202pp.

Sivakesava, S. and Irudayaraj, J. 2001. A Rapid spectroscopic technique for determining honey adulteration with corn syrup. *Journal of Food Science.* (66)6:787-791.

Snedaker, S. C., Glynn, P. W., Rumbold, D. G., and Corcoran, E. F. 1995. Distribution of *n*-alkanes in marine samples from southeast Florida. *Marine Pollution Bulletin* 30(1):83-89.

Stone, M. 1974. Cross-validatory choice and assessment of statistical predictions (with discussion). *Journal of the Royal Statistical Society B*, 36(2):111-147.

Syunyaev, Z. R., R. M. Balabin, M. R., I. S. Akhatov, S. I., and Safieva, O. J. 2009. Adsorption of Petroleum Asphaltene onto Reservoir Rock Sands Studied by Near-Infrared (NIR) Spectroscopy, *Energy & Fuels*, 23(3):1230-1236.

Szolar, O. H. J., Rost, H., Braun, R., and Loibner, A. P. 2002. Analysis of polycyclic aromatic hydrocarbons in soil. Minimizing sample pre-treatment using automated Soxhlet with ethyl acetate as extraction solvent. *Analytical Chemistry*, 74:2379-2385.

Tate, R. 2002. Soil microbiology. 2nd edition. New York, John Wiley & Sons. 508 p.

Thurston, T. J., Brereton R. G., Foord D. J., and Escott R. E. A. 2004. A principal components plots for exploratory investigation of reactions using ultraviolet-visible spectroscopy: application to the formation of benzophenone phenylhydrazone. *Talanta* 63:757-769.

Tisdale, L.S., Nelson, W. L., Beaton, J. D., and Havlin, J. L. 1993. Soil and fertilizer nitrogen. *In* Soil fertility and fertilizers. Fifth edition. p109-175.

Torres R. 2003. El papel de los microorganismos en la biodegradación de compuestos tóxicos. *Ecosistemas*, 12(2):1-5.

Triola, F. M. 2006. Estadística. 9ª edición. Pearson Addison Wesley. México. 838p.

United States Department of Agriculture (USDA). 1996. Soil survey laboratory methods manual. Soil Survey Investigations. Washington D.C, USA, 693p.

US EPA (Environmental Protection Agency). 1996. Method 3540C. Soxhlet extraction. United States Environmental Protection Agency, SW-846 Manual. Washington, DC: Government Printing Office. 1-8p.

US EPA (Environmental Protection Agency). 1996. Method 3560. Supercritical fluid extraction of total recoverable petroleum hydrocarbons. United States Environmental Protection Agency, SW-846 Manual. Washington, DC: Government Printing Office. 1-8 p.

US EPA (Environmental Protection Agency). 2007. Method 3535A. Solid-phase extraction (SPE). United States Environmental Protection Agency, SW-846 Manual. Washington, DC: Government Printing Office. 1-23 p.

US EPA (Environmental Protection Agency). 2007. Method 3545A. Pressurized fluid extraction (PFE). United States Environmental Protection Agency, SW-846 Manual. Washington, DC: Government Printing Office. 1-16 p.

US EPA (Environmental Protection Agency). 2007. Method 3550C. Ultrasonic extraction. United States Environmental Protection Agency, SW-846 Manual. Washington, DC: Government Printing Office. 1-17 p.

Valcárcel, M. and Cárdenas, S. M. 2000. Automatización y miniaturización en química analítica. Springer – Verlag Ibérica S. A. Barcelona, España.

Van Hamme, J. D., Singh, A., and Ward, P. O. 2000. Recent advances in petroleum Microbiology. *Microbiol. Mol. Biol. Rev.* 67(4): 503-549.

Vásquez, D. R., Abadia, B., Arreaza, L. C. 2004. Aplicación de la espectroscopía de reflectancia en el infrarrojo cercano (NIRS) para la caracterización nutricional del pasto Guinea y del grano del maíz. *CORPOICA*, 5(1):49-55.

Viñas, C. M. 2005. Biorremediación de suelos contaminados por hidrocarburos: caracterización microbiológica, química y ecotoxicológica. Tesis de doctorado en biología. Universidad de Barcelona. España. Pp. 352.

Wang, Y., Veltkamp, D. J., and Kowalski, B. R. 1991. Multivariate Instrument Standardization. *Analytical Chem.* 63:2750.

Wang, Z. and Fingas, M. F. 2003. Development of oil hydrocarbon fingerprint and identification techniques. *Marine Pollution Bulletin*, 47:423-452.

Wang, Z., Dean, T., and Kowalski, B. R. 1995. Additive background correction in multivariate instrument standardization. *Analytical Chemistry*, 67:2379-2385.

Williams, P. C. 1987. Variables affecting near infrared reflectance spectroscopic analysis. In: Williams, Ph., and Norris, K. Near infrared technology in the agricultural and food industries. American Association of Cereal Chemists, Minnesota, USA.

Williams, P. C. 2003. Near-infrared technology-getting the best out of the light. A short course in the practical implementation of near infrared spectroscopy for user, 5.3th edition. Nanaimo, Canada: PDK Projects Inc.

Williams, P. C., and Sobering, D. C. 1993. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *J. Near Infrared Spectrosc.* (1):25-32.

WinISI III. 2005. Infracsoft International, Copyright© by LLC. Version 1.60.

Wold, S., and Sjöström, M. 1977. SIMCA: A Method for analyzing chemical data in terms of similarity and analogy. Research Group for Chemometrics, Institute of Chemistry, Umeå University.

Wu, Y., Junfeng, Ji., Ch, J., Gong, P., Liao, Q., Tian, Q., and Ma, H. 2007. A mechanism study of reflectance spectroscopy for investigating heavy metals in soils. *Soil Sci. Soc. Am. J.* 71(3):918-926

Zavala, C. J. 2009. Estudio integral de la cuenca hidrológica del río Tonalá y lagunas

del Carmen y Machona. Informe no publicado. H. Cárdenas, Tabasco.

Zavala, C. J. F., Gavi, R. R. H., Adams, S. H. R., Ferrera-Cerrato, D. J., Palma-López, D. J., Vaquera, H. Y., y Domínguez, E. J. M. 2005. Derrames de petróleo en suelos y adaptación de pastos tropicales en el activo Cinco Presidentes, Tabasco. *Terra*. 23:293-302.

Zavala, C. J., Gavi, R. F., Adams, S. R. H., Ferrera-Cerrato, R., y Palma-López, D. J. 2007. Hidrocarburos del petróleo y tecnología de biorremediación para suelos de Tabasco. *In: Plan de uso sustentable de los suelos de Tabasco. Volumen II.* 2007. Palma-López, D. J., y Triano, S. A. (eds.). Colegio de Postgraduados, ISPROTAB, Gobierno del Estado de Tabasco, México. 180p.

Zavala, C. J., Palma-López, D. J., Ortíz, S. C. A., Gutiérrez, C. Ma. del C., y Ruíz, B. A. 2003. Utilidad de las clases de tierras campesinas en suelos de aluvión reciente. *In: impacto ambiental en las tierras del campo petrolero Samaria, Tabasco.* 2003. Zavala, C. J., Gutiérrez, C. Ma. del C., Palma-López, D. J. (eds.). Colegio de Postgraduados, CONACYT, CCYTET. Villahermosa, Tabasco, México. 131p.

Zhang, L., Small, C. W., and Arnold, M. A. 2002. Calibration standardization algorithm for partial least-squares regression: application to the determination of physiological levels of glucose by near-infrared spectroscopy. *Analytical Chemistry* 74:4097-4108.

Zwanziger, W. H., and Förster, H. 1998. Near infrared spectroscopy of fuel contaminated sand and soil. I. preliminary results and calibration study. *Journal of Near Infrared Spectroscopy*, 6(1):189-197.