



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Química e Ingeniería Química

Unidad de Posgrado

**Predicción de la presión de vapor de esteres ftálicos
empleados como plastificantes en función de su
estructura molecular por descriptores moleculares**

TESIS

Para optar el Grado Académico de Magíster en Química con
mención en Fisicoquímica

AUTOR

Víctor Andrés ACUÑA RUIZ

ASESOR

Scila REÁTEGUI SÁNCHEZ

Lima, Perú

2018

ACTA DE TITULACION



Universidad Nacional Mayor de San Marcos
 Universidad del Perú. Decana de América
FACULTAD DE QUÍMICA E INGENIERÍA QUÍMICA
VICEDECANATO DE INVESTIGACIÓN Y POSGRADO
 Calle Germán Amezaga N° 375 - Ciudad Universitaria, Teléfono 619-7000 anexo 1204 - Lima 1 - Perú
 "Año del Diálogo y la Reconciliación Nacional"
 "Decenio de la Igualdad de Oportunidades para mujeres y hombres"

UNIDAD DE POSGRADO ACTA DE SUSTENTACIÓN DE TESIS DE GRADO ACADÉMICO DE MAGÍSTER


A los seis días del mes de noviembre del año dos mil dieciocho, a las 10:30 horas en el auditorio CERSEU de la Facultad de Química e Ingeniería Química, se reunió el Jurado de Tesis presidido por el Dr. Mario Ceroni Galloso, el Dr. Eduardo Montoya Rossi, el Dr. Marcos A. Loroño, el Mg. Holger Maldonado García y la Mg. Scila Reátegui Sánchez, para calificar la sustentación de la tesis titulada: **PREDICCIÓN DE LA PRESIÓN DE VAPOR DE ESTERES FTÁLICOS EMPLEADOS COMO PLASTIFICANTES EN FUNCIÓN DE SU ESTRUCTURA MOLECULAR POR DESCRIPTORES MOLECULARES**, presentado por el Quim. Víctor Andrés Acuña Ruiz para optar el Grado Académico de Magíster en Química con mención en Físicoquímica.

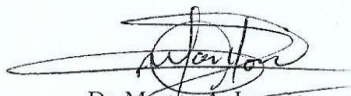
Hecha la exposición y absueltas las preguntas formuladas por el jurado, éste acordó la siguiente calificación de acuerdo a lo establecido por el Art. 61 del Reglamento General de Estudios de Posgrado, aprobado por R.R. N° 00301-R-09 del 22 de enero del 2009:

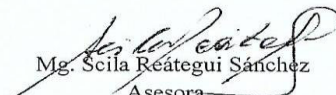
(17) DIECISIETE - muy BUENO

Habiendo sido aprobada la sustentación de la tesis, el Jurado recomendó que la Facultad proponga que se le otorgue el Grado Académico de Magíster en Química con mención en Físicoquímica, al Quim. Víctor Andrés Acuña Ruiz.


El acto académico de sustentación concluyó a las: 12:00 horas.


 Mg. Holger Maldonado García
 Miembro


 Dr. Marcos A. Loroño
 Miembro


 Mg. Scila Reátegui Sánchez
 Asesora


 Dr. Eduardo Montoya Rossi
 Miembro


 Dr. Mario Ceroni Galloso
 Presidente

DEDICATORIA

Dedicado a la memoria de mi madre, a mi esposa Liz por su invaluable apoyo sin el cual me hubiera sido imposible el desarrollo de ésta Tesis y también a mis hijas Andrea, Victoria e hijo Víctor

AGRADECIMIENTO

Mi agradecimiento a la Profesora Magister Scila Reátegui Sánchez por su asesoramiento científico a la presente Tesis

Mi reconocimiento y especial agradecimiento al Doctor PhD. Juan Ladislao Arroyo Cuyubamba por su sugerencia y apoyo del tema de Tesis desarrollada

INDICE GENERAL

Carátula	i
Acta de Titulación	ii
Dedicatoria	iii
Agradecimientos	iv
Indice General	v
Lista de Tablas	viii
Lista de Figuras	xi
Resumen	xiii
Abstract	xvi
CAPITULO 1: INTRODUCCION	1
1.1 Situación Problemática	1
1.2 Formulación del Problema	6
1.3 Justificación Teórica	6
1.4 Justificación Práctica	8
1.5 Objetivos	8
1.5.1 Objetivo General	8
1.5.2 Objetivo Específico	9
CAPITULO 2: MARCO TEORICO	10
2.1 Marco Filosófico o Epistemológico de la Investigación	10
2.2 Antecedentes de la Investigación	12
2.3 Bases Teóricas	14
2.3.1 Relación Cuantitativa Estructura Propiedad	14
2.3.2 Descriptor molecular	15

2.3.3 Características del Descriptor Molecular	15
2.3.4 Clasificación de los Descriptores Moleculares	16
2.3.5 Descriptor Molecular RDF o Código de Distribución Radial	19
2.3.6 E-DRAGON 1.0 software	22
2.3.7 SMILES (Simplified Molecular Input Line Entry System)	23
2.3.8 CORINA software (COoRdINAtes)	24
2.3.9 BuildQSAR software	25
2.3.10 Regresión Lineal Múltiple (RLM)	26
2.3.11 Reporte QSPR	36
2.3.12 Validación del Modelo	36
CAPITULO 3: METODOLOGIA	40
3.1 Recolección de Esteres Ftálicos y selección de sus valores de presión de vapor	40
3.2 Cálculo de los Descriptores Moleculares	47
3.3 Determinación del modelo	49
CAPITULO 4: RESULTADOS y DISCUSION	53
4.1 Conjunto de Esteres DMP, DAP, DBP, BBP y DEP	53
4.2 Conjunto de Esteres DNOP, 610P, DIOP, DEHP, BOP y DHP	59
4.3 Descriptores moleculares determinados	65
CONCLUSIONES	78
REFERENCIAS BIBLIOGRAFICAS	79
ANEXOS	85
Anexo 1a: Comunicación VCCLAB	85
Anexo 1b: Comunicación Kode - Solutions	85

Anexo 2: Ejemplo de Cálculo de Descriptores moleculares RDF.

Ftalato DEHP	86
Anexo 3: Descriptor RDFv	108
Anexo 4: Descriptor RDFe	109
Anexo 5: Descriptor RDFu	110
Anexo 6: Descriptor RDFm	111

LISTA DE TABLAS

Tabla 1. Clasificación de los Descriptores de acuerdo a la dimensión de su Representación Molecular Producción anual de trabajos QSPR-QSAR	17
Tabla 2. Interfase BuildQSAR	26
Tabla 3. Abreviatura y Nombre de los Esteres Ftálicos	40
Tabla 4. Valores de Presión de Vapor de Ésteres Ftálicos	41
Tabla 5. Presiones de Vapor, Masa molar y fórmula global de Ésteres Ftálicos	43
Tabla 6. Identificación de los CAS y SMILES de los ftalatos	44
Tabla 7. Descriptores RDFu proporcionados por E-Dragon 1.0.Grupo 1	47
Tabla 8. Valores descriptores RDF calculados por E-Dragon 1.0	48
Tabla 9. BuildQSAR. DMP, DAP, DBP, BBP	49
Tabla 10. BuildQSAR. DOP, 610P, DIOP, BOP, DEHP	50
Tabla 11. Coeficientes asociados a cada descriptor, a través de sus parámetros estadísticos R, F, s y Q ² reportados por BuildQSAR	51
Tabla 12 . Modelo de regresión lineal de variable múltiple. DMP, DAP, DBP, BBP	52
Tabla 13. Modelo de regresión lineal de variable múltiple. DOP, 610P, DIOP, BOP, DEHP	52

Tabla 14. Descriptores RDF115v, RDF060e	53
Tabla 15. . Análisis de varianza 1	54
Tabla 16. Coeficiente de correlación 1	55
Tabla 17 Parámetros del modelo 1	55
Tabla 18. Capacidad predictora 1	56
Tabla 19. Residuales1 para los DMP, DAP, DBP y BBP	56
Tabla 20. Presión de vapor observada Vs calculada DMP-DAP- DBP-BBP-DEP	58
Tabla 21. Descriptores RDFu, RDFm	59
Tabla 22. Análisis de varianza 2	60
Tabla 23. Coeficiente de correlación 2	60
Tabla 24. Parámetros del modelo 2	61
Tabla 25. Capacidad predictora 2	61
Tabla 26. Residuales 2: DNOP, 610P, DIOP, DEHP, BOP	62
Tabla 27. Presión de vapor observada Vs. Calculada DNOP, 610P, DIOP, DEHP, BOP	64
Tabla 28. Resultado General. Presión de vapor observada Vs calculada.	65
Tabla 29. Contribución porcentual jerárquica por átomo. Ftalatos: DAP, DBP, BBP, DMP, DEP, - RDF060e	67

Tabla 30. Contribución porcentual jerárquica por átomo. Ftalatos: DAP, DBP, BBP, DMP, DEP - RDF115v	68
Tabla 31. Contribución porcentual jerárquica por átomo. Ftalatos: DOP, 610P, DIOP, DEHP, BOP, DHP - RDF065u	69
Tabla 32. Contribución porcentual jerárquica por átomo. Ftalatos: DOP, 610P, DIOP, DEHP, BOP, DHP - RDF095u	70
Tabla 33. Contribución porcentual jerárquica por átomo. Ftalatos: DOP, 610P, DIOP, DEHP, BOP, DHP - RDF145m	71
Tabla 34. Principales átomos contribuyentes. Ftalatos: DAP, DBP, BBP, DMP, DEP	72
Tabla 35. Principales átomos contribuyentes. Ftalatos: DNOP, 610P, DIOP, DEHP, BOP, DHP	72

LISTA DE FIGURAS

Figura 1. Histograma de la producción anual de trabajos QSPR-QSAR	7
Figura 2. Histograma, por área temática, de trabajos QSPR y QSAR	7
Figura 3. Descriptor de Función de Distribución Radial–RDF	20
Figura 4. Interfase E-Dragon 1.0	23
Figura 5. Métrica F. Regla de decisión de rechazo hipótesis nula	33
Figura 6. Homocedasticidad. Varianza residual.	35
Figura 7. Fórmulas desarrolladas y estructura de Ftalatos	44
Figura 8. Formato de texto de salida imprimible del software E-Dragon 1.0	47
Figura 9. Espectro descriptor RDFv de los ésteres DMP-DEP-DAP- DBP-BBP	54
Figura 10. Espectro descriptor RDFe de los ésteres DMP-DEP-DAP- DBP-BBP	54
Figura 11. Homocedasticidad Residuales 1	57
Figura 12. Normalidad Residuales 1	57
Figura 13. Independencia Residuales 1	57
Figura 14. Valor Observado Vs. Calculado de los DMP-DEP-DAP- DBP-BBP	58
Figura 15. Espectro del descriptor RDFu DNOP-610P-DIOP-BOP- DEHP-DHP	59
Figura 16. Espectro del descriptor RDFm. DNOP-610P-DIOP-BOP- DEHP-DHP	60
Figura 17. Homocedasticidad. Residuales 2	62
Figura 18. Normalidad. Residuales 2	63

Figura 19. Independencia. Residuales 2	63
Figura 20. Valor observado Vs. calculado de los ésteres DNOP-610P- DIOP-BOP-DEHP-DHP	64
Figura 21. Valor observado Vs. calculado de los ésteres DMP-DEP-DAP- DBP-BBP- DNOP-610P-DIOP-BOP-DEHP-DHP	65
Figura 22. Estructura mínima común. Ftalatos DMP, DAP, DBP, BBP y DEP	73
Figura 23. Estructura mínima común. Ftalatos DNOP- 610P- DIOP- DEHP- BOP- DHP	73
Figura 24. Modelo contributivo estructural posicional aproximado. Ftalatos DMP-DAP- DBP- BBP-DEP	74
Figura 25. Modelo contributivo estructural posicional aproximado. Ftalatos DNOP- 610P- DIOP- DEHP- BOP-DHP	74

RESUMEN

La presión de vapor de los ésteres ftálicos fueron calculados por el uso de descriptores moleculares derivados de sus estructuras. Los ésteres ftálicos fueron tomados del artículo de Staples et. al. 1997. El software E-Dragon 1.0 fue empleado para calcular los descriptores moleculares. El software BuilQSAR fue usado para calcular el modelo de Regresión Lineal Multiparametro. Como resultado se obtuvo que los descriptores moleculares RDF (Función de Distribución Radial) son los que mejor calculan la presión de vapor. Los valores presión de vapor calculados fueron de gran concordancia con los valores observados. La capacidad predictiva del modelo se probó con el coeficiente de correlación de validación cruzada Q^2 .

Palabras clave: QSPR, Descriptores RDF, presión de vapor, ésteres ftálicos, MLR, Regresión Lineal Multivariable . E-Dragon 1.0. BuildQSAR.

ABSTRACT

Vapor pressure calculation of phthalic esters used as plasticizers have been calculated by using chemical molecular descriptors derived from their structure. The phthalic esters were taken from Staples *et al.* 1997. E-Dragon 1.0 software was used to calculate the molecular descriptors. BuilQSAR software was also used to calculate the model MLR (Multiparameter Linear Regression). As a result RDF (Radial Distribution Function) molecular descriptors were determined as the better molecular descriptors to compute the vapor pressure. The values of vapor pressure calculated were in good agreement with the observed values. The predictive capacity of the model was demonstrated by the significant cross validated correlation coefficient Q^2 .

Key words: QSPR, RDF descriptors, vapor pressure, phthalic esters, MLR, Multivariable Linear Regression. E-Dragon 1.0. BuildQSAR.

CAPITULO1: INTRODUCCION

1.1 Situación Problemática

Los ésteres **ftálicos** son una familia de sustancias químicas sintéticas, de bajo peso molecular relacionada estructuralmente con el ácido orgánico conocido como ácido ftálico. Debido a sus bajos puntos de fusión y elevados puntos de ebullición los ésteres ftálicos se utilizan como plastificantes, sobre todo se añade al policloruro de vinilo (PVC), como fluidos de transferencia de calor y portadores de carga másica (Staples, Peterson, Parkerton y Adams, 1997). Material polimérico; como plásticos, pinturas o adhesivos, mejoran su flexibilidad cuando se les adiciona estas sustancias (Wypych, 2004). Los plastificantes son usados en pisos vinílicos, pesticidas, fragancias, repelente de insectos, cables, mangueras, tapicería, películas, pinturas, adhesivos, tintas y diversos productos de policloruro de polivinilo (Mersiowsky, Weller y Jorgen, 2001).

Para el año 2006, la producción anual de los ésteres ftálicos excedía los 5 millones de toneladas (Mackintosh, Maldonado, Ikonomou y Gobas, 2006) y cada año se desechan al medio ambiente toneladas de material plástico PVC, en cuya composición se encuentran los ftalatos, hasta en un 93%, como plastificantes mayoritarios (Mersiowsky, Weller y Jorgen, 2001).

El efecto de los ésteres ftálicos sobre el medio ambiente y la salud han sido estudiados por diversos organismos los que dan a conocer acerca de su acción ambiental y en la salud. Se ha determinado que los ftalatos y sus metabolitos son potencialmente perjudiciales para el humano y el ambiente, debido a su toxicidad hepática, teratogenicidad y características cancerígenas (Matsumoto, Hirata y Ema, 2008). Desde 1991, la Environmental Protection Agency (EPA) viene realizando estudios sobre sus efectos en el ambiente y estableció máximos de concentración para el Di-etil ftalato (DEP), difundido en el Reglamento Nacional de Agua Potable Primaria, US EPA, 1991. Posteriormente; en base a consideraciones de toxicidad, su prevalencia en el ambiente, su uso generalizado y exposición a los humanos, la EPA (Plan de Acción para ftalatos, US EPA, 2012)

extendió su plan de control ambiental a ocho ésteres ftálicos: dibutil ftalato (DBP), diisobutil ftalato (DIBP), butil benzil ftalato (BBP), di-n-pentil ftalato (DnPP), di-2-etilhexil ftalato (DEHP), di-n-octil ftalato (DnOP), diisononil ftalato (DINP) y diisodecil ftalato (DIDP).

Debido a que los ésteres ftálicos no están covalentemente unidos al polímero del cual forma parte, favorece su capacidad de migrar a la superficie de la matriz y desde allí pueden perderse o diseminarse mediante una variedad de procesos físicos (Stanley, Robillard, Staples, 2003). Luego de migrar a la superficie de la matriz polimérica y dependiendo de la fase a la cual se transfiere, los ésteres ftálicos pueden difundirse en el medio ambiente de varios modos.

Los procesos de liberación de los ésteres ftálicos y otros compuestos al medio ambiente, denominados destino ambiental, los encontramos descritos en la Base de Datos de Sustancias Peligrosas (HSDB) de la Biblioteca Nacional de Medicina de EEUU (US National Library of Medicine) y en ella están referidos tres formas de diseminación o destino ambiental para el di-n-octil ftalato.

Basándose en los valores de 4.38-6.3 para el $\log K_{oc}$ (Schuurmann, 2006 y Staples, Peterson, Parkerton y Adams, 1997), la HSBD presenta el destino ambiental terrestre citando al esquema de clasificación propuesto por Swann et. al. respecto de la movilidad química de los ftalatos en suelos (Swann, Laskowski, McCall, 1983)[9]. De acuerdo a dicho valor debe esperarse que el DnOP se mantenga inmóvil en el suelo. Considerando los valores de las constantes de la ley de Henry, basadas en su presión de vapor de 1.0×10^{-7} mm Hg (Staples, Peterson, Parkerton y Adams, 1997) [3] y su solubilidad en el agua de 0.022 mg/L (Defoe, 1990), es posible prever que la volatilización del DnOP a partir de la superficie de suelos húmedos sea un importante proceso de diseminación o de destino; por lo que se espera que la adsorción al suelo atenúe la volatilización y basado en su presión de vapor se espera que el DnOP no se volatilice de la superficie de suelos secos.

Para la diseminación acuática, la HSDB citando a Swann (Swann *et al.*, 1983) y considerando los valores de $\log K_{oc}$ de 4.38-6.3 (Schuurmann, 2006; y Staples, Peterson, Parkerton, Adams, 1997) se espera que el DnOP se adsorba a sólidos suspendidos y sedimentos. Así mismo, considerando el valor estimado de 2.6×10^{-6} atm-m³/moles para la constante de la ley de Henry; valor derivado de su presión de vapor de 1×10^{-7} mm Hg (Staples, Peterson, Parkerton y Adams, 1997) y su solubilidad en agua de 0.022 mg/L (Defoe, 1990), se espera la volatilización de las superficies de agua. El destino o diseminación atmosférica de compuestos orgánicos semivolátiles (Blideman, 1988) es considerada por la HSDB de acuerdo al modelo de partición gas/partícula. Considerando el valor de 1.0×10^{-7} mmHg a 25°C que tiene la presión de vapor del DnOP, se espera que en el ambiente atmosférico se encuentre tanto en las fases de vapor y de partícula (Staples, Peterson, Parkerton y Adams, 1997). Los radicales hidroxilos, producidos fotoquímicamente en la atmósfera, degradan la fase vapor del DnOP y usando un método de valoración estructural (Interfase de Programa de Estimación EPI, US EPA, 2015) se estima un tiempo de vida media de 19 h, calculado a partir de su constante de velocidad de 2.1×10^{-11} cm³/molécula-s a 25°C. La fase individual de DnOP puede ser retirada del aire mediante deposiciones húmedas o secas y se le identifica por la existencia de una banda de absorción intensa a 274.5 nm la que se extiende más allá de los 290 nm (Sadler Index, 1994) lo cual le hace susceptible de sufrir fotólisis directa de la luz solar.

El efecto de los plastificantes sobre la salud está determinada por la concentración del plastificante en el medio ambiente y el cálculo para la determinación de la cantidad de sustancia liberada al medio ambiente se basa, particularmente, en tres coeficientes del proceso de destino ambiental: Coeficiente de partición octanol-agua (K_{ow}), volatilidad y las constantes de la ley de Henry. Así mismo, el valor de la volatilidad está determinado por la presión de vapor del plastificante en estudio y los valores de la presión de vapor del plastificante usado en la estimación del

destino ambiental pueden ser encontrados en Handbooks y artículos científicos. Sin embargo, los valores de la presión de vapor informados no son unívocos para cada especie química, sino que en algunos casos son contradictorios y en otros existe ausencia de datos. De la información encontrada, no todos los datos corresponden a valores obtenidos experimentalmente sino que muchos de ellos han sido determinados mediante cálculos, desarrollados a partir de ecuaciones como la de Clausius o Antoine. Las contradicciones obedecen a la diversidad de métodos o a su uso en condiciones extremas, quedando a criterio del investigador la elección del valor que considere más “apropiado”; generándose una incertidumbre en el resultado de la estimación ambiental que ha empleado un valor de presión de vapor, obtenido bajo dichas condiciones. Tal situación está expresada en el Handbook o Manual de Química de Lange (Lange’s Handbook, 1999) donde manifiestan que: “El editor difícilmente puede tener el conocimiento experto de todas las ramas de la ciencia ni la intuición necesaria para decidir en todos los casos qué valor en particular deberá publicar, especialmente cuando muchos diferentes valores, para la misma constante, están dados en la literatura”.

Al respecto, similar opinión se encuentran en el Handbook o Manual de propiedades físico químicas y destino ambiental de compuestos orgánicos de Mackay y colaboradores (Mackay, Wan, Kuo, y Sum, 2006) cuando indican que: “La tarea de investigar el destino o diseminación química local, regional o global es complicada debido al gran número (en incremento) de compuestos químicos potencialmente implicados, de la incertidumbre de sus propiedades físico-químicas y por la ausencia del conocimiento de las condiciones ambientales prevalentes: temperatura, pH y velocidad de deposición de materia sólida de la atmosfera al agua o del agua hacia sedimentos profundos. Más aún, en propiedades como la solubilidad donde los valores reportados están frecuentemente en conflicto. Algunas de ellas son medidas en forma precisa, otras aproximadamente y algunas son estimadas por esquemas de correlación diferentes, de sus estructuras

moleculares. En algunos casos, unidades o identidades químicas son reportadas erróneamente. El usuario de tales datos tiene, entonces, la difícil tarea de seleccionar el “mejor” o “correcto” valor. Existe por ello una justificable preocupación de que las deducciones resultantes acerca del destino o diseminación ambiental podrían tener errores sustanciales. Por ejemplo, el potencial de evaporación podría ser muy sub estimado si se selecciona o elije un valor pequeño para la presión de vapor.

Esta divergencia de valores es confirmada en trabajos de investigación como la de Cousins y Mackay (Cousins y Mackay, 2000) en la que expresan: “ ... la inspección de estos datos revelan que los valores reportados de propiedades físico-químicas específicas pueden variar en diferentes órdenes de magnitud... Las fuentes de tales imprecisiones no siempre son conocidas pero las posibles causas de variación incluyen las diferencias en la metodología usada para medir la propiedad, diferencias en métodos analíticos para determinar las concentraciones, la naturaleza extrema de algunas mediciones (esto es cerca al límite de detección del método analítico empleado) y también errores y omisiones de los investigadores”.

En relación a los diversos valores reportados sobre las presiones de vapor de los ftalatos, Tienpont (Tienpont, 2004) indica: “En el caso de los ésteres ftálicos, los cuales exhiben solo volatilidad limitada, la medida directa de la presión a elevadas temperaturas combinada con la extrapolación a temperatura ambiental usando las ecuaciones de Clausius-Clapeyron o de Antoine resultan en importantes imprecisiones, para medidas de presiones de vapor...”

Rohac, Ruzicka y colaboradores (Rohac, Ruzicka K, Ruzicka V, Zaitsan, Kabo, Diky y Aim, 2004) coinciden sobre la inconsistencia de datos cuando expresan: “A pesar que se encuentran disponibles en la literatura datos para la presión de vapor, en general son mutuamente inconsistentes y sujetos a grandes incertidumbres”.

1.2 Formulación del Problema

Durante la determinación de la presión de vapor de una misma especie química, aplicando métodos exploratorios y/o experimentales se observan valores disímiles. ¿Será posible calcular la presión de vapor de los ésteres ftálicos empleados como plastificantes, aplicando métodos de cálculo que relacionan la estructura con la actividad o propiedad?

1.3 Justificación Teórica

Actualmente, el cálculo de las propiedades físicas y químicas, desde el punto de vista de la estructura de las especies químicas, se ha constituido en un vasto y activo campo de investigación; introducido como alternativa a las determinaciones que se realizaban en condiciones experimentales límites con resultados inciertos o contradictorios, obtenidos mediante la extrapolación de las ecuaciones Clausius o de Antoine, desarrolladas matemáticamente.

Un estudio bibliométrico, basado en las citas bibliográficas de artículos publicados en el 2014 por Li, Jianxin Hu, and Yuh-Shan Ho acerca de la producción mundial de trabajos publicados entre 1993 y 2012 para el cálculo de propiedades, físicas, químicas y biológicas basados en la estructura de la molécula, reporta un promedio de 900 artículos por año. El incremento de publicaciones de 164 en 1993 a 713 en 2006 registra un crecimiento de cuatro veces, siendo la Revista de Información y Modelaje Químico (Journal of Chemical Information and Modeling) la más prolífica en publicaciones.

Las innovaciones metodológicas en la construcción de modelos moleculares y modelaje incentivan la producción de la investigación en el campo del diseño y síntesis de fármacos y química informática; mientras que las validaciones externas y la confiabilidad incentivan la investigación en el campo de las ciencias ambientales. La predicción de punto final, algoritmos estadísticos, descriptores moleculares, fueron establecidos como

campos de intensa investigación. Los artículos provenientes de los países desarrollados fueron largamente mayores en las citas bibliográficas, mientras que las de los países en desarrollo fueron los de mayor crecimiento.

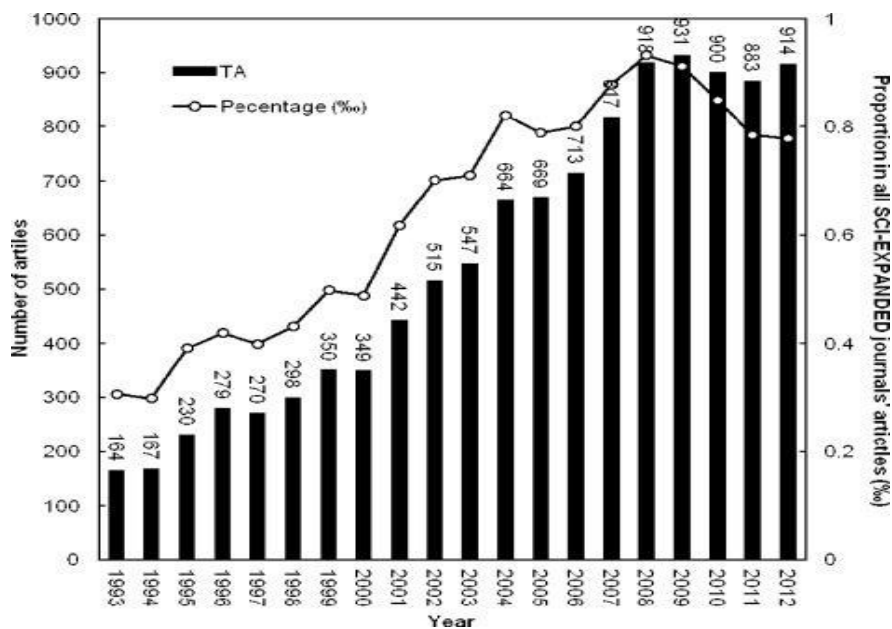


Fig. N°1. Histograma de la producción anual de trabajos QSPR-QSAR.

(Fuente: Li Li, et al, 2014)

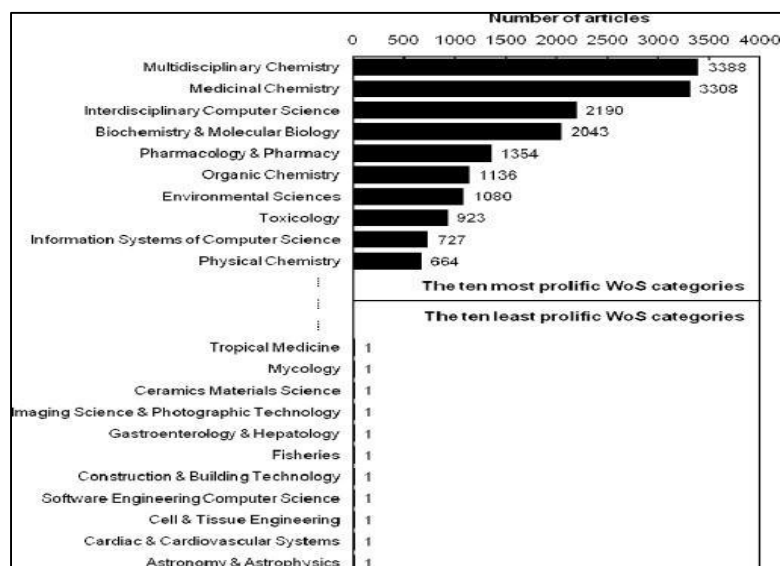


Fig. N°2. Histograma, por área temática, de trabajos QSPR y QSAR.

(Fuente: Li Li, et al. 2014)

1.4 Justificación Práctica

Como propiedad físico-química, la presión de vapor juega un rol importante en la prevención de incendios, explosiones, diseño, control de procesos y optimización de procesos de ingeniería (Lawson, 1980, Redeker, 1997; Sandler, Lin, & Sum, 2002; Stoll, 2005). Los datos de valores de presión de vapor también son usados para estimar la viscosidad de líquidos, las entalpías de evaporación, coeficientes de partición aire-agua y otras propiedades importantes físico-químicas de las diferentes especies químicas.

El valor de la presión de vapor posee amplia aplicación práctica en diversos campos. Se considera como de importancia significativa en la determinación de la distribución de los contaminantes entre la atmósfera y el suelo, en la identificación de áreas de destino ambiental a través del transporte y el reparto de dichos compuestos en el agua, aire y suelo. Así mismo, es posible usar datos de la presión de vapor de químicos tóxicos para estimar la velocidad de evaporación después de un derrame.

Sin embargo, la carencia de recursos e instalaciones hace que la determinación de la presión de vapor, en el laboratorio, de un creciente número de productos químicos sea impracticable (Katritzky *et al.*), dando paso a la Relación Cuantitativa Estructura Propiedad (QSPR) de la presión de vapor con los descriptores teóricos. Su investigación tiene un valor significativo, no requiere datos experimentales y encuentra aplicación en los métodos de modelamiento molecular basados en la técnica de relación estructura-propiedad/actividad (QSPR/QSAR) (Lium Huangm & Wang, 2003; Sandlres *et al.*, 2002; Stoll, 2005)

1.5 Objetivos

1.5.1 Objetivo General

Calcular las presiones de vapor de los ésteres ftálicos, empleados como plastificantes, tomando como base a descriptores moleculares de Función de Distribución Radial (RDF).

1.5.2 Objetivo Específico

Identificar los tipos de descriptores moleculares RDF que calculan la presión de vapor de los ésteres ftálicos empleados como plastificantes.

CAPITULO 2: MARCO TEORICO

2.1 Marco Filosófico o Epistemológico de la investigación

La epistemología aborda los fundamentos de la ciencia, su relación con otras ramas del saber y el tipo de conocimiento que produce. Immanuel Kant, a través de su obra **“Crítica de la Razón Pura y Fundamentos Metafísicos de la Ciencia Natural”** señalaba que la química nunca podría ser una ciencia porque su conocimiento no podía ser matematizado. Sin embargo, al publicar Lavosier el Tratado Elemental de Química incorpora, de manera sistémica, el tratamiento matemático a sus observaciones. Siendo la materia, desde el punto de vista físico y/o químico, susceptible de observación y considerando que la ciencia física emplea intensamente la matemática, una forma de afianzar el pensamiento kantiano se dio mediante la aplicación de los métodos físicos a la química. Esto derivó en la físico-química, la que se enfocó en el estudio de la producción de calor y energía de las reacciones químicas; dando lugar a la termoquímica, la electroquímica y a la química cuántica. La noción de matematizar, exhaustiva y casi excluyentemente, la observación química condujo a Paul Dirac a afirmar que **“. . . las leyes necesarias para la teoría matemática, de gran parte de la física y la totalidad de la química, son bien conocidas. La dificultad radica en que la aplicación exacta de estas leyes nos conduce a ecuaciones demasiadas complicadas para ser resueltas. Es deseable desarrollar métodos aproximados de aplicación de la mecánica cuántica, que faciliten la explicación de aspectos fundamentales de los sistemas atómicos complejos, sin recurrir excesivamente a la computación”** (Pflughaupt, 2009) y (Villaveces, 2000). La aplicación de métodos de aproximación y procesos computacionales extraordinariamente rápidos; como los procesos *ab initio*, semi empíricos y la teoría de la densidad funcional, aún no han podido resolver, de manera práctica, el aspecto predictivo de los

problemas de la química. Luego, es posible pensar si la aplicación de la matemática al fenómeno químico es el único camino para presentar explicaciones causales aceptables, considerando que aún los procedimientos exclusivamente matemáticos también aceptan aproximaciones y simplificaciones.

Así mismo, el enfoque matemático del fenómeno químico no excluyó pero si relegó a un segundo plano otros enfoques diferentes; como aquéllos derivados de la estructura de la especie química. Uno de ellos tiene su antecedente notable en la predicción del peso atómico de los elementos, propuesto por Mendeleiev en su tabla periódica. Un enfoque moderno que busca predecir las propiedades de las especies químicas se basan en los trabajos de Hans y Fujita. Ellos orientaron su trabajo a la predicción de la actividad biológica en base a la estructura química. Esta relación, entre la estructura y propiedad cuantificada, fue denominada Relación Cuantitativa Estructura Propiedad (QSAR en inglés) y el éxito logrado le permitió expandirse a otras áreas, como las referidas a las propiedades físicas y/o químicas, conocidas como Relación Cuantitativa Estructura Propiedad (QSPR en inglés),

Sin embargo, fueron los trabajos de Hammett (Hammett L., 1937) y Hansch y Fujita (Hansch C. y Fujita T., 1964) los pioneros en demostrar que las propiedades de los compuestos químicos se corresponden o pueden ser explicados en función de su estructura y delinearon los pasos fundamentales para establecer una relación cuantitativa entre una característica estructural y una determinada propiedad. De este modo, se inició el desarrollo de trabajos basados en el método propuesto por Hammett y Hansch, denominándose QSAR al modelo predictivo de actividad biológica basado en la estructura del compuesto (Quantitative Structure Activity Relationship) y QSPR al modelo predictivo para determinar una propiedad química o física basado en la estructura del compuesto (Quantitative Structure Property Relationship).

2.2 Antecedentes de la investigación

Trabajos sobre QSAR y QSPR, relacionados a ftalatos y otras sustancias orgánicas, se encuentran publicados en los últimos años del siglo pasado y con más frecuencia en este siglo. Los autores son mencionados en orden cronológico.

Katritzky, *et al*, 1995, refiere que cualquier regularidad encontrada se puede utilizar para racionalizar el comportamiento de moléculas en el sistema y pronosticar las propiedades de otros compuestos que pertenecen a un grupo dado; es decir “buscar regularidad en la forma en la cual varían las propiedades moleculares y cómo esas variaciones dependen de la estructura molecular”. “Más aún, las relaciones encontradas entre estructura y propiedad (o actividad biológica) puede ser importante en el desarrollo de nuevas teorías, las cuales a su vez podrían explicar tanto el fenómeno observado como elucidar los mecanismos de fenómenos físicos y químicos o de actividad biológica”

En 1997, Liang y Gallagher presentan un modelo QSPR para estimar la presión de vapor de compuestos con diferentes grupos funcionales orgánicos sin individualizar sus resultados. Adicionalmente, toman como grupo de cotejo datos promediados o extrapolados y no aquéllos que sean exclusivamente experimentales.

En 1998, Katritzky, Wang, Sild, Tamm y Karelson, propusieron 411 compuestos para ser estudiados. Entre ellos se encuentran hidrocarburos, hidrocarburos halogenados y compuestos que contienen oxígeno y nitrógeno. Sin embargo, en el trabajo publicado no se indican claramente los grupos funcionales oxigenados con los que se realizó el estudio. Así mismo, no se individualiza o identifica a los componentes del estudio.

En 1999, Thomsen, Rasmussen y Carlsen describen, por las relaciones estructura-actividad (SAR) y relación cuantitativa estructura-propiedad (QSPR), algunas propiedades de los ftalatos como la solubilidad, los

coeficientes de partición agua–octanol y el coeficiente de sorción de suelos.

En 1999, Sacan, Ozkul y Erdem, presentan modelos QSPR para solubilidad del n-octanol/agua (S), el coeficiente de partición del n-octanol/agua (K_{ow}), la constante de la ley de Henry (H) para las dibenzo-p-dioxinas policloradas (PCDDs), los dibenzo-p-furanos (PCDFs) y ftalatos basándose en parámetros topológicos, descriptores moleculares semi empíricos y el momento dipolar.

En el año 2000, para hallar la presión de vapor, Cousins y Mackay sugieren aplicar la ecuación que obtuvieron para calcular los coeficientes de partición de ésteres ftálicos a partir de las solubilidades en las interfaces aire/agua (K_{AW}), octanol/agua (K_{OW}) y octanol/aire (K_{OA}). Sin embargo, los resultados reportados para las presiones de vapor difieren en más de dos órdenes de magnitud, en varios casos, respecto de los resultados experimentales.

En el 2005, se reportaron trabajos QSPR como el de Subhash y Mills, para las presiones de vapor de los ésteres ftálicos; pero no muestran los resultados ni identifican cada una de las especies químicas estudiadas.

En el año 2007, en una ampliación de este trabajo, realizado por Katritzky, Slavov, Dobchev y Karelson, proponen un modelo para predecir la presión de vapor de compuestos orgánicos, manteniendo la falta de identificación individual de los compuestos evaluados.

En el 2009, Chunhui Lu basándose en su índice propuesto y aplicando modelos QSPR, estudió las propiedades ambientales de los ftalatos; entre ellas su solubilidad en medios acuosos. Sin embargo, no indica la fuente de adquisición de los datos de valores experimentales de comparación y en el caso de la presión de vapor solo presenta 5 valores denominados experimentales. Extiende sus resultados a 53 tipos de ftalatos, no obstante que sólo presenta valores aceptables para tres datos experimentales.

En el 2013, Pramanik y Roy, presentan modelos de destino ambiental (QSFR) de 455 compuestos orgánicos; aunque no especifican a ningún compuesto en particular, y

En el 2015, Abraham y Acree, presentan trabajos QSPR sobre diversas propiedades de los ésteres ftálicos, relacionados a sus propiedades bioquímicas.

2.3 Bases teóricas

Una molécula se considera como la mínima agrupación de átomos, de una determinada sustancia, con una estructura característica que se manifiesta en sus propiedades físico químicas, una de las cuales es su presión de vapor. Atkins, 2006, define a esta propiedad como la presión que ejerce el vapor de una sustancia, en un volumen cerrado, en equilibrio con su fase condensada. Establecer la correlación o una relación apropiada entre la estructura de la molécula y una propiedad determinada, puede resultar útil para predecir propiedades no medidas y aún hasta desconocidas. La relación cuantitativa estructura-propiedad (QSPR) relaciona la estructura química molecular con una amplia variedad de sus propiedades físicas, químicas, ambientales, biológicas; incluyendo biomédicas, toxicológicas y ecotoxicológicas o tecnológicas como la concentración crítica de micela, temperatura de transición vítrea, cinética de vulcanización de cauchos, etc.

2.3.1 Relación Cuantitativa Estructura Propiedad

La Relación Cuantitativa Estructura-Actividad (QSAR) es un método cuantitativo que busca establecer un modelo que relacione, mediante descriptores, las características estructurales de compuestos químicos con su actividad biológica. El modelo de actividad biológica fue extendido a otras propiedades físicoquímicas, denominado Relación Cuantitativa Estructura-Propiedad (QSPR). Diferentes propiedades o comportamientos de moléculas químicas en el campo de las Relaciones Cuantitativas (QR): Estructura-Reactividad (QSRRs), Estructura-

Cromatografía (QSCRs), Estructura-Toxicidad (QSTRs), Estructura-Electroquímica (QSERs) y Estructura-Biodegradabilidad (QBRs), han sido investigados por S. Yousefinejad, B. Hemmateenejad, 2015.

2.3.2 Descriptor Molecular

Todeschini *et al.*, 2000, define al descriptor de una estructura molecular como el resultado final de un proceso lógico-matemático que transforma la información química codificada en un número útil, a la representación simbólica de una molécula o de algún experimento estandarizado. Según Gasteiger y Engel, 2003, es la representación matemática de una molécula resultado de un procedimiento que transforma la información estructural codificada de la forma de una molécula. Para permitir la construcción de modelos con métodos estáticos y redes neuronales artificiales, la representación matemática debe ser independiente del tamaño y número de átomos de la molécula. La información contenida en un descriptor de estructura depende de dos factores principales: a. La representación molecular del compuesto y b. El algoritmo usado para el cálculo del descriptor.

Del mismo modo Faulon y Bender, 2010, definen al descriptor molecular como una propiedad abstracta, en la mayoría de los casos numérica, de una estructura molecular derivada de algún algoritmo que describe un aspecto específico de la molécula.

2.3.3 Características del Descriptor Molecular

Un descriptor molecular útil, simple, interpretable, etc. debe reunir algunos requerimientos matemáticos. Las propiedades básicas que un descriptor molecular en particular, según Todeschini, *et al.*, 2000, debe cumplir son:

1. Los descriptores que emplean numeración de átomos en su definición, deberán usar una única forma canónica de numeración basada en reglas inequívocas. Será invariante con respecto a la identificación y numerado de los átomos de la molécula.
- 2.- Invarianza respecto a la roto-traslación de la molécula; es decir el valor de un descriptor molecular no dependerá de los valores absolutos de las coordenadas numéricas que definen las posiciones de los átomos en relación a algún origen arbitrario; es decir los valores de un descriptor no deben ser diferentes al considerar la posición de la molécula respecto de un eje fijo de referencia.
- 3.- Una definición algorítmicamente computable, no ambigua o inequívoca se refiere a una expresión matemática computable donde los términos no deben ser ambiguos y estar claramente visibles en la molécula.
- 4.- Valores en un rango numérico adecuado para el conjunto de moléculas donde será aplicado.

2.3.4 Clasificación de los Descriptores Moleculares

En la actualidad y basados en diferentes criterios, son muchas las formas en que pueden ser clasificados los descriptores moleculares. Gasteiger & Engel, 2003, presentan una clasificación como la que se muestra en la Tabla N°1.

Tabla N° 1. Clasificación de los Descriptores de acuerdo a la dimensión de su Representación Molecular.

Representación Molecular	Descriptor	Ejemplos
0 D	Contador de átomos, contador de enlaces, peso molecular, suma de propiedades atómicas.	Peso molecular, peso molecular promedio, número de átomos, número de enlaces, número de enlaces múltiples, número de enlaces aromáticos, número de enlaces rotacionales, suma de volúmenes atómicos van der waals
1 D	Contador de fragmentos	Número de carbonos primarios, secundarios, terciarios, Número de átomos donantes de enlaces H, número de átomos aceptores de enlaces H
2 D	Descriptores topológicos	Índice Zagreb, índice Wiener, índice Balaban J, índice de conectividad chi (X), índice de conectividad kappa (k), cuenta de caminos moleculares
3 D	Descriptores geométricos	Excentricidad molecular. Radio de giro, descriptor 3D MoRSE, función de distribución radial (código RDF), descriptores WHIM, descriptor GETAWAY, vector 3D autocorrelación.
4 D		Coordenadas 3D + muestreo de conformaciones

(Fuente: adaptada a partir de Johann Gasteiger y Engel, 2003).

Una clasificación alternativa no excluyente y proporcionada por Katritzky *et al*, 1995, plantea la tipificación de la molécula en función de su: composición, parámetros moleculares clásicos, estructura electrostática, características tridimensionales y los considerados como modificados.

Descriptores Constitucionales: Considera fundamental la composición de la molécula más que su topología, geometría o estructura electrónica. Estos descriptores son simples y no deben ser obviados en los estudios QSAR/QSPR.

Descriptores Topológicos: Probablemente son los descriptores más usados e incluyen parámetros moleculares clásicos conocidos como

índice de: Wiener, Randic y conectividad molecular Kier & Hall. Estos tipos de descriptores obviamente son más sensibles a la topología molecular (i.e. conectividad molecular) y en particular a la ramificación de la molécula. Algunos descriptores topológicos reflejan la composición molecular, aunque en menor extensión que los constitucionales.

Descriptores Electroestáticos: Reflejan la distribución de carga parcial o la electronegatividad de los átomos, aunque también en muchos casos estén referidos a la topología molecular. Los descriptores de la “familia de áreas superficiales parcialmente cargadas”, en proporciones comparables, son el reflejo de la electrostática, la geometría y las características topológicas de una molécula. La distribución de cargas parciales puede ser calculada, independiente del programa, por uno o más procedimientos no empíricos dentro del programa QSAR/QSPR o cualquier método deseado.

Descriptores Geométricos: En descriptores químico-cuánticos ab-initio y otros semi-empíricos son importantes el tamaño y la forma de la molécula. Estos han incrementado su uso extensamente, son rutinarios y cada vez más disponibles. Dichos cálculos pueden proporcionar vasta y variada información sobre la estructura química, incluyendo datos electrostáticos y de geometría. Lo más importante de tales cálculos es que pueden proveer información acerca de las propiedades electrónicas internas de la molécula, no disponibles por otros medios. Con lo que se extienden los descriptores químico-cuánticos a las áreas de aplicación de las técnicas QSAR/QSPR. Entre ellos, los más frecuentemente empleados son la energía del orbital molecular más alto ocupado (HOMO) y la energía del orbital molecular más bajo desocupado (LUMO), la distribución de carga poblacional de Mulliken y los momentos dipolares.

Descriptores Modificados: Para obtener un valor “promedio” del descriptor, la mayor parte de estos descriptores pueden ser normalizados al dividirlos entre el número de átomos. Mediante los Programas

avanzados de QSAR/QSPR es posible calcular los valores standard del descriptor y también algunas modificaciones de ellos (Katritzky *et al.*, 1995).

2.3.5 Descriptor Molecular RDF o de Función de Distribución Radial

El descriptor molecular RDF tiene sus antecedentes en los códigos 3D-MoRSE, muy valioso en la conservación de las características moleculares y poco probable de ser interpretado directamente. Todeschini *et al.*, 2000, afirma que “La función de distribución radial reúne todos los requerimientos para un descriptor 3D. Es independiente del número de átomos y el tamaño de la molécula; es única considerando el arreglo tridimensional de los átomos y es invariante respecto de la traslación y rotación de la molécula entera. Adicionalmente, el código RDF puede estar referido a átomos específicos o rangos de distancias para representar información específica en cierta estructura espacial 3D, así como describir el impedimento estérico de la estructura y/o propiedades o actividad de la molécula”. Así mismo, “El código RDF es interpretable mediante reglas simples y por lo tanto proporciona la posibilidad de convertir de nuevo el código en la estructura”. Gasteiger y Engel, 2003, al investigar tipos de descriptores asociados y al incorporar la idea de la función de distribución radial, utilizan la dispersión de rayos X en sus investigaciones, describen la función de distribución radial y transforman propiedades a descriptor molecular. La función de distribución radial, conocida en fisicoquímica y física, es posible aplicarla a un modelo molecular 3D porque describe la distribución de distancias de puntos en dicho espacio y se obtiene al incluir en la ecuación de la función de distribución radial (2.1), propiedades atómicas las cuales establecerán las características de los átomos en el ambiente molecular. La función de distribución radial, $g(r)$, es la probabilidad de las distancias atómicas entre los átomos i y j , de los N átomos del espacio 3D y usualmente se calcula mediante un número discreto de intervalos.

$$g(r) = \sum_i^{N-1} \sum_{j>i}^N e^{-B(r_i-r_{ij})^2} \quad (2.1)$$

El término exponencial contiene al parámetro de resolución B, se interpreta como factor térmico que determina el movimiento de los átomos y determina el ancho de los picos individuales en la distribución de probabilidad.

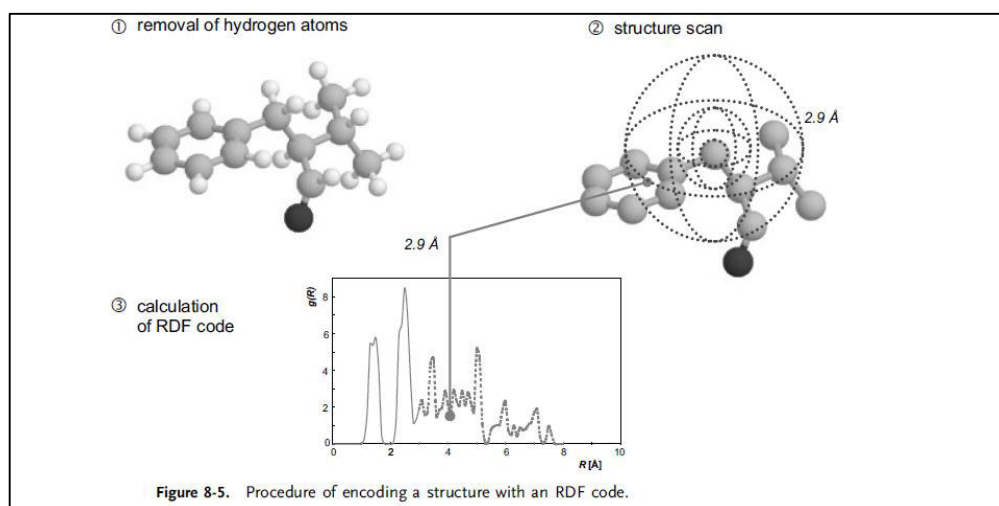


Fig.Nº3. Descriptor de Función de Distribución Radial–RDF.

(Fuente: Gasteiger y Engel, 2003)

Hemmer, M., 2007, plantea una ligera modificación a la forma general de la RDF y conduce a un descriptor molecular, denominado Código de Distribución de Función Radial. Este incluye propiedades atómicas, las cuales condicionarán las propiedades o rasgos atómicos característicos del entorno molecular, y Gasteiger, J. y Engel, T., 2007, afirman que las funciones deben ser normalizadas.

$$g(r) = \sum_i^{N-1} \sum_{j>i}^N A_i A_j e^{-B(r_i-r_{ij})^2} \quad (2.2)$$

$$g(r) = f \sum_i^{N-1} \sum_{j>i}^N A_i A_j e^{-B(r_i-r_{ij})^2} \quad (2.3)$$

A, representa a las propiedades atómicas características de los átomos i y j. N es el número de átomos y f es un factor de normalización o escalamiento. E-Dragon 1 escala o normaliza a las funciones de distribución radial en función a las propiedades del átomo de carbono (comunicación privada entre Kode Solutions desarrolladores de E-Dragon 1.0 y el autor. Ver Anexo 1). El término exponencial contiene la distancia r_{ij} entre los átomos i y j. El parámetro de arreglo B, define la distribución de probabilidad de las distancias individuales. $g(x)$ se calcula para un número discreto de puntos con intervalos definidos. El código RDF puede ser usado en diferentes tareas y puede ajustarse a los requisitos de la información a la que representa.

Las propiedades atómicas utilizadas en la ecuación (2.2) permiten discriminar cualquier propiedad atribuible a un átomo. Hemmer, Steinhauer, Gasteiger, 1999, refieren que la función de distribución permite acceder a información acerca de las distancias interatómicas de toda la molécula y la oportunidad de acceder a información valiosa como tipos de átomos, distancias de enlace, tipos de anillos, tipos de sistemas planares y no planares. Este hecho es una consideración muy importante y valiosa para una elucidación de código asistida por computadora. De esta forma, la función de distribución radial reúne los requerimientos considerados, en particular la invarianza respecto de las traslaciones lineales.

Existen varias formas de generar descriptores basados en el concepto de la función de distribución radial. Todas las funciones RDF presentadas anteriormente se pueden calcular incluyendo las propiedades A de átomos individuales. Estas propiedades pueden ser insertadas, como un producto, en un término pre exponencial y da lugar al descriptor de RDF

ponderado por propiedades, denominado en publicaciones anteriores como código RDF. Este método transforma la dimensión de la frecuencia en una dimensión de frecuencia ponderada por la propiedad. Son dos tipos de propiedades atómicas empleadas: 1. Propiedades atómicas estáticas: las que son características para un tipo de átomo pero que son independientes de la molécula individual por ejemplo número atómico, volumen atómico, potencial de ionización y en el caso más simple los números atómicos que proveen información acerca del tipo de átomos relacionados con una cierta distancia. 2. Propiedades atómicas dinámicas: dependen del entorno químico del átomo y es característico de la molécula, entre ellos la carga atómica parcial, la polarizabilidad atómica y la electronegatividad parcial.

Hemmer M., 2007, afirma que, con la finalidad de representar la información parcial o completa de un ambiente químico particular, los descriptores RDF pueden ser restringidos a tipos de átomos específicos o rangos de distancia; pudiendo describir el impedimento estérico de una reacción, o la relación estructura/actividad de una molécula. Finalmente, los descriptores RDF pueden ser interpretados usando reglas simples debido a la posibilidad de convertir el vector en la correspondiente estructura 3D.

2.3.6 E-DRAGON 1.0 Software

Virtual Computational Chemistry Laboratory (VCCLAB) ha desarrollado el software E-Dragon 1.0 para calcular descriptores moleculares con formato de entrada de estructura molecular en SMILES (Mauri, et. Al., 2006), disponible en <http://www.vcclab.org/lab/edragon/>. Los descriptores incluidos en el software son Descriptores Constitucionales (48), Contadores de Paso y Ruta (47), Índices Informativos (47), Índices de Adyacencia de Frontera (107), Índices de Carga Topológica (21), Perfiles Moleculares de Randic (41), Descriptores RDF (150), Descriptores WHIM (99), Contadores de Grupos Funcionales (154), Descriptores de Carga (14), Descriptores Topológicos (119), Índices de

Conectividad (33), Autocorrelaciones 2D (96), Autovalores o Valores Propios de Carga (64), Índices basados en valores propios o Autovalores, Descriptores Geométricos (74), Descriptores #D MoRSE (160), Descriptores GETAWAY, (197), Fragmentos Centrados en Átomos (120), Propiedades Moleculares (31) . Los descriptores RDF proporcionados por E-DRAGON 1.0 son los siguientes

1. RDFu no ponderado.
2. RDFm ponderado por la masa atómica.
3. RDFv ponderado por el volumen Van der Waals.
4. RDFe ponderado por electronegatividad de Sanderson.
5. RDFp ponderado por polarizabilidad.

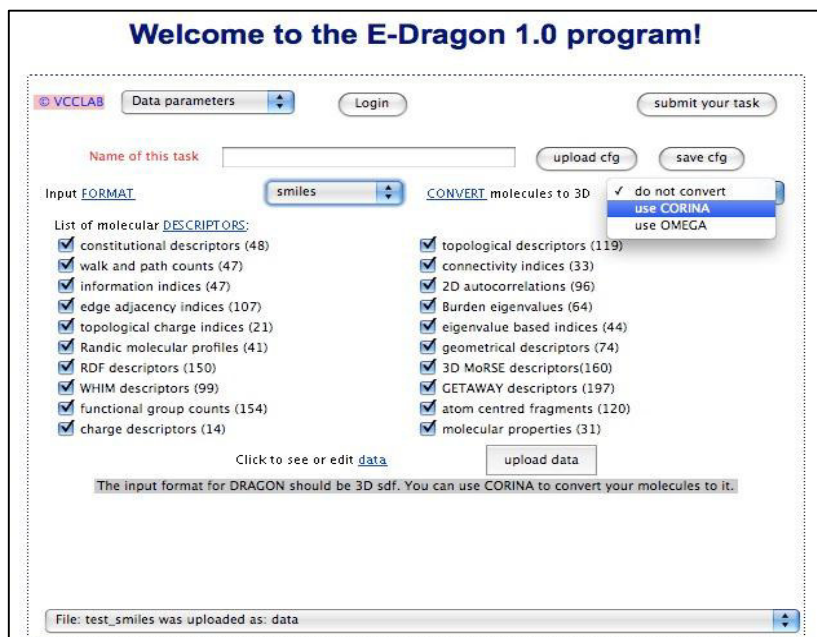


Fig. N°4. Interfase E-Dragon 1.0. (<http://www.vcclab.org/lab/edragon/>)

2.3.7 SMILES (Simplified Molecular Input Line Entry System)

SMILES es un lenguaje químico comprensible y simple en el que las moléculas y reacciones se especifican usando los caracteres de ASCII que representan símbolos del átomo y del enlace. El desarrollo de

SMILES fue iniciada por David Weininger (Weininger D., 1988) utilizando el concepto de un gráfico con nodos como átomos y bordes como enlaces para representar una molécula, los paréntesis se usan para indicar los puntos de ramificación y las etiquetas numéricas designan los puntos de conexión del anillo. SMILES convierte la estructura química 3D en una cadena de símbolos que son fácilmente entendidos por softwares de computadoras. SMILES es usado para ingresar estructuras químicas en programas de estimación de parámetros físicoquímicos de incidencia ambiental como EPI Suite (Estimation Program Interface Suite) y ECOSAR (Ecological Structure Activity Relationship) que son softwares de propiedad de la Agencia de Protección Ambiental de los Estados Unidos de Norteamérica (U.S. EPA). SMILES tiene ventajas sobre las tablas de conexión y su forma canónica es empleada como identificador universal de una estructura química específica y correcta de una molécula o reacción química, a partir de una cadena SMILES. Su gramática básica también incluye información isotópica, configuración de enlaces dobles y quiralidad, todo lo cual conduce al denominado isómero SMILES. Es una marca registrada de Daylight Chemical Information Systems, Inc., con información disponible en <http://www.daylight.com/smiles/index.html>. También existen diversos programas que transforman las estructuras químicas en notación SMILES.

2.3.8 CORINA software (COoRdINAtes)

El CORINA software es definida por Gasteiger, Rudolph y Sadowski, 1990, como un convertidor automático a estructuras 3D. Además se tiene entre otros software a COBRA, CHEMX, MOLGEO, CORINA, CONVERTER, CONCORD y ALCOGEN.

Hemmer, et.al., 1999, y Levine I., 2001, refieren que este software trabaja con datos introducidos con SMILES, considerando longitudes de enlace y ángulos estándar, tabulados al descomponer la estructura en una parte

acíclica y otra en estructuras de anillos de más de 9 átomos y otra en anillos de hasta 9 átomos. También combina fragmentos monocéntricos con ángulos y longitudes de enlace estándar, generando coordenadas internas y ángulos diedros para construir una molécula. CORINA es aplicable a toda la química orgánica y en ella se procesan casi todas las estructuras que puedan ser expresadas en una notación de enlace valencia con información estereoquímica. Esto se demuestra en la conversión de grandes bases de datos, como la del Instituto Nacional del Cáncer de los Estados Unidos de Norteamérica (NCI) que ha desarrollado alrededor de 250,000 estructuras. No existe límite superior respecto al tamaño o anillos de las moléculas y se permite estructuras hasta con seis átomos vecinos. Para optimizar la geometría de cada anillo en la molécula, CORINA usa un pseudocampo de fuerzas tipo mecánica molecular simplificado y por defecto genera una conformación de baja energía por cada estructura ingresada.

Vracko y Gasteiger, 2002, afirman que se ha demostrado que el programa CORINA provee parámetros geométricos que generan modelos comprobables, en calidad, con los parámetros geométricos obtenidos al optimizar procedimientos químicos cuánticos; siendo su alta velocidad un aspecto relevante de este software. Para el estudio de una molécula en particular, usando un CPU SUN IPX, fueron necesarios 10 minutos; mientras que al usar el procedimiento AM1 y CORINA sólo requirió un décimo de segundo.

2.3.9 BuildQSAR Software

Daniel Barbosa de Oliveira y Anderson Coser Gaudio, dos profesores de la Universidade Federal do Espirito Santo de Brasil (de Oliveira y Gaudio, 2000) desarrollaron el BuildQSAR software. Este es de análisis estadístico desarrollado para la construcción y análisis cuantitativo de modelos por medio del análisis regresional, aplicable a propiedades

fisicoquímicas y biológicas. Disponible en <http://www.profanderson.net/files/buildqsar.php>

Tabla N°2. Interfase BuildQSAR. (Fuente: BuildQSAR software)



ID	Structure	Comment	Y1	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19
1	DMP		266.6448	4.798	5.504	4.085	15.008	3.23	5.582	5.624	8.001	12.99	2.092	8.355	5.583	7.985	2.878	1.38	0.859	0	0	0
2	DAP		21.331584	7.073	7.547	7.668	20.128	7.274	9.379	10.233	6.856	18.971	3.595	15.262	11.931	6.812	8.56	7.202	9.437	3.154	5.061	3.5
3	DBP		3.5997048	10.137	10.99	5.605	43.951	17.434	7.081	13.371	11.07	33.164	2.39	12.474	15.397	10.019	20.642	8.479	15.898	11.329	8.555	13
4	BBP		0.666612	9.659	10.733	6.173	38.98	10.343	6.697	12.209	15.062	27.14	8.05	13.274	14.648	14.666	12.614	15.226	17.004	12.418	10.494	12



ID	Structure	Comment	Y1	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19
1	DNOP		0.01333224	17.254	18.305	8.655	83.51	39.737	7.081	18.98	12.551	65.805	2.743	15.25	45.324	9.142	57.281	8.473	22.986	27.046	11.981	
2	610P		0.65327976	17.254	18.305	8.655	83.508	39.729	7.081	18.978	12.533	65.843	2.741	15.237	45.326	9.147	57.29	8.477	20.769	31.008	12.922	
3	DIOP		0.1333224	17.254	18.306	9.352	79.93	41.337	15.437	23.324	17.433	57.952	8.365	22.264	35.265	10.774	44.252	9.757	23.886	20.618	14.16	
4	BOP		0.014665464	13.696	14.648	8.124	59.823	29.616	11.677	24.486	17.572	46.044	9.941	21.625	30.915	18.479	31.703	14.245	26.754	19.477	13.628	
5	DEHP		0.01333224	17.254	18.306	9.743	78.854	41.767	20.328	28.92	27.873	53.181	19.2	31.589	48.156	28.621	31.108	23.644	40.699	29.959	30.296	

2.3.10 Regresión Lineal Múltiple (RLM)

El fundamento matemático de los modelos QSAR/QSPR se sustenta en el principio de la polilinealidad (PPL). En ella una continua y singular dependencia entre la propiedad P_i , experimentalmente medible, y algún factor estructural intrínseco de la molécula X_j es asumida como lineal en un cierto dominio de este factor (X_j). Dependiendo de la forma funcional de la dependencia $P_i(x_j)$, en la vecindad del punto x_j , la suposición de linealidad puede ser válida o inválida, Katritzky, Lobanov, Karelson, 1995. En el caso que el PPL sea válido, la propiedad experimental podría depender solamente de un solo factor estructural, X_i , y la ecuación lineal correspondiente de parámetro único puede ser encontrada usando la ecuación (2.4) a través del método de los mínimos cuadrados lineales. Alternativamente, la propiedad podría depender linealmente de varios factores estructurales X_j , y la correspondiente regresión de parámetro

múltiple puede ser hallada aplicando la ecuación (2.5) a través del método de mínimos cuadrados multilineal.

$$P_i = a_{0i} + a_{1i}X_1 \quad (2.4)$$

$$P_i = a_{0i} + a_{1i}X_1 + a_{2i}X_2 + a_{3i}X_3 + \quad (2.5)$$

En el caso que la aproximación del PPL no sea válida, algunas formas de ecuaciones no lineales deben ser aplicadas para describir la dependencia de la propiedad experimental respecto del factor estructural. Si se usa una transformación no lineal del factor estructural, tal como las potencias cuadráticas, logaritmos o términos cruzados (ecuación 2.6) entonces la dependencia funcional formal de P_i de X'_j continúa manteniéndose lineal y la correspondiente ecuación de regresión puede ser hallada usando el mismo método de mínimos cuadrados multilineal (ec 2.6).

$$(cuadrado)X'_j = X_j^2 \text{ o } (\logaritmo)X'_j = \ln X_j \text{ o } (\text{término-cruzado})X'_j = X_j X_k \quad (2.6)$$

$$P_i = a'_{0i} + a'_{1i}X'_1 + a'_{2i}X'_2 + a'_{3i}X'_3 + \quad (2.7)$$

Los métodos basados en algoritmos de mínimos cuadrados parciales y de regresión lineal múltiple no son los únicos métodos, también existen métodos de análisis como el Proceso Gaussiano, el análisis discriminante lineal, la regresión logística, la red neural de regresión general, la red neural probabilística, etc.; pero los métodos de mínimos cuadrados parciales y los de regresión lineal múltiple son los más usados y reportados en publicaciones de revistas especializadas como PubMed y Scopus.

Históricamente, la Regresión Lineal Múltiple (RLM) es el primer método de calibración multivariable usado en el diseño de modelos. Su simplicidad y la facilidad para interpretarlos, han permitido una extensa aplicación a casi todas las ramas de los estudios QSAR y QSPR. En la RLM, se establece una relación lineal entre las características

moleculares de una molécula, expresada generalmente como un vector x , a un descriptor y su actividad o propiedad. La forma general de RLM es:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_kx_k + e \quad (2.8)$$

y : Variable dependiente de las variables independientes x_1, x_2, \dots, x_k

x_1, \dots, x_k : Variables independientes o descriptores moleculares.

b_1 a b_k : Coeficientes del descriptor x_1 a x_k , b_0 es la constante o intercepto del modelo y ; el residuo de la actividad / propiedad, que no está cubierto por el modelo, es identificado por e . Este término, también conocido como error aleatorio, es la parte del modelo que las variables independientes o predictores no explican y contiene el efecto, sobre y , de todas las variables diferentes de x_1, x_2, \dots, x_k .

La notación matricial de esta ecuación, para un grupo de moléculas, es:

$$Y = Xb + e \quad (2.9)$$

La magnitud de los coeficientes del modelo son valores normalizados que denotan la importancia relativa de los descriptores y sus signos indican el tipo de contribución, positiva o negativa, a la actividad o propiedad.

La solución de mínimos cuadrados para la estimación de b podría ser definido como:

$$b = (C^T C)^{-1} C^T y \quad (2.10)$$

Por su parte, Yousefinejad y Hemmateenejad, 2015, indican que el principal inconveniente de RLM es su limitación para modelar descriptores con alto grado de colinealidad. Esto podría conducir a modelos con coeficientes de regresión imprecisos [358]. Además, el número de variables en el modelo no deberán exceder al número de

moléculas en el grupo de ensayo. Desde el punto de vista matemático, el número de moléculas debería ser, al menos, igual al número de descriptores más uno.

La aceptación del modelo y por consiguiente la dependencia de la variable estudiada en función de las variables predictores implica que b_1, b_2, \dots, b_k sean diferentes de cero. Como la ecuación del modelo se obtiene a partir de muestras, su aceptación es probabilística y la validez dentro de un rango de probabilidad indica que el modelo no es producto del azar. El rango generalmente aceptado es del 95% por lo que el enunciado de la hipótesis nula es que no existe una relación lineal entre las variables independientes o predictores. La hipótesis alterna, al negar la hipótesis nula, afirma que existe una relación lineal entre variables predictores y la variable dependiente u observada. La decisión de aceptación del modelo se da por el análisis de la varianza, comúnmente denominada ANOVA, realizado en función a las métricas estadísticas F y de significancia en un 95% de probabilidad. Valores altos de F, asociados con una significancia menor de 0.05, permiten rechazar la hipótesis nula y aceptar la hipótesis alterna; es decir permite aceptar el modelo expresado a través de la ecuación de regresión lineal multivariable.

Para asegurar su aplicabilidad y que sea satisfactorio, el modelo de regresión debe reunir las condiciones de significancia y predictibilidad. Además, en el modelo las variables independientes o descriptores moleculares deben ser ortogonales; es decir no deben ser colineales entre sí o estar correlacionados, a fin de tener la certeza a qué descriptor corresponde la desviación de la variable observada (presión de vapor). Los Descriptores completamente independientes tienen un coeficiente de correlación de 0.0 y se denominan ortogonales. Casos ideales son difíciles de encontrar y la correlación de dos descriptores normalmente no debería ser mayor de 0.6. Gasteiger y Engel, 2003, reportan

coeficientes de correlación aceptables entre descriptores, en rangos menores de 0,4 a 0,9.

La bondad de ajuste o calidad de una RLM es usualmente evaluada por R^2 denominado coeficiente de correlación cuadrado. Éste correlaciona los valores experimentales de la variable observada, en nuestro caso la presión de vapor, y las variables seleccionadas como predictores o explicativas de los valores de la presión de vapor, los descriptores RDF. R^2 tiene un valor entre cero y uno que indica la proporción de la variación en la variable dependiente que esta explicada por la ecuación del modelo de regresión. Un modelo es considerado aceptable cuando los valores de R^2 son cercanos a la unidad. En el documento Guía para validar los modelos de la Relación Cuantitativa Estructura Actividad QSAR (2007), la Organización para la Cooperación y el Desarrollo Económico (OCDE), define a R^2 por:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.11)$$

Donde:

$$0 \leq R^2 \leq 1 \text{ y } R = \sqrt{R^2}.$$

y_i : valor observado, \hat{y}_i : valor de la variable calculada.

\bar{y} : valor promedio de los valores observados

$$SS_{\text{Res}} \text{ (Suma de residuales cuadrados)} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.12)$$

$$SS_{\text{T}} \text{ (Suma total de cuadrados)} = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (2.13)$$

Los modelos de regresión también emplean el Coeficiente de correlación cuadrado ajustado R^2_{adj} . En los modelos de regresión lineal multivariable a medida que se incrementan el número de variables explicativas el valor de R^2 también aumenta, pudiendo llegar a una falsa situación en la que el coeficiente de R^2 se aproxima a la unidad debido al exceso de variables explicativas y no al buen ajuste o calidad del modelo. En consideración de esta situación se emplea el término R^2_{adj} , el cual es una métrica estadística variante de R^2 con la notable diferencia de que solo aumenta en valor cuando se adiciona una variable independiente que mejora al modelo y disminuye si se adiciona alguna variable que no lo mejora. Montgomery, D., Peck, E., Vining G. (2012), definen a R^2_{adj} como

$$R^2_{adj} = 1 - \frac{SSRes / (n-p)}{SST / (n-1)} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n-p)}{\sum_{i=1}^n (y_i - \bar{y})^2 / (n-1)} \quad (2.14)$$

Donde:

n: Número de observaciones o datos.

p: Número de variables explicativas o de regresión, más el intercepto.

La utilidad del modelo se determina con la métrica o estadígrafo F, el cual depende de la relación entre la varianza explicada por el modelo respecto a la varianza residual. Evalúa a la hipótesis nula y refiere que todos los coeficientes de la regresión son nulos o iguales a cero; es decir que el modelo no tiene significancia estadística. A mayor valor de F mayor utilidad del modelo; mientras que si el valor es cero el modelo no posee significancia o utilidad. Kutner M., Nachtsheim J., Neter J., Li W. (2005), lo definen como:

$$F = \frac{MSR}{MSE} \quad (2.15)$$

Donde:

MSR: Son los cuadrados medios del modelo.

MSE: Son los cuadrados medios del error.

$$MSR = \frac{SSR}{p-1} \quad (2.16)$$

$$MSE = \frac{SSE}{n-p} \quad (2.17)$$

Donde:

$(p - 1)$: Grados de libertad del modelo.

p : Número de predictores +1.

$(n - 1)$: Grados de libertad del error.

n : Número de observaciones.

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (2.18)$$

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.19)$$

SSR: Suma de cuadrados del modelo de regresión.

SSE: Suma de cuadrados del error.

A partir de lo cual se tiene:

$$F = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / (p - 1)}{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - p)} \quad (2.20)$$

Donde:

y_i : Valor observado.

\hat{y}_i : Valor calculado o predicho.

\bar{y} : Media de los valores observados.

n : Número de datos u observaciones.
 p: Número de predictores +1

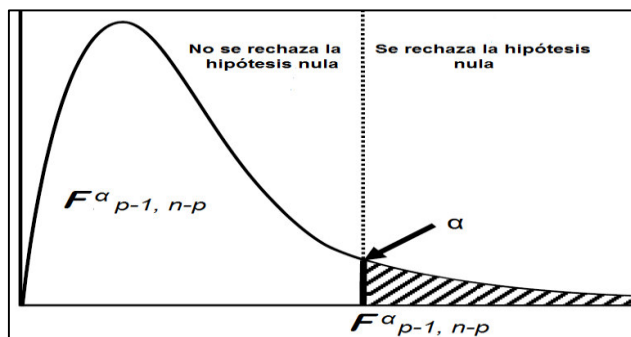


Fig.Nº5. Métrica F para la decisión de rechazo de la hipótesis nula.

(Fuente: propia)

La decisión de rechazo de la hipótesis nula se obtiene al comparar el valor de la métrica F, obtenida del análisis de la varianza ANOVA, con el valor de F estadístico, obtenido de la tabla de distribución normal que corresponde a $F^{\alpha}_{p-1, n-p}$ y al extrapolar a la ec. 2.20; α es 0.05, (p-1) son los grados de libertad del numerador y (n-p) son los grados de libertad del denominador. Si el valor de F estadístico es menor que el de ANOVA, se rechaza la hipótesis nula y se acepta la hipótesis alterna mediante la cual se afirma que existe relación significativa entre la variable dependiente y las variables independientes o predictores del modelo.

Como la predictibilidad es el atributo del modelo obtenido al calcular la propiedad investigada y obteniéndose valores para especies no consideradas en la construcción del modelo cercanos a la experimental, ésta se determina por la técnica denominada validación cruzada mediante el procedimiento "uno a uno", del cual se obtienen los estadígrafos o métricas PRESSp y Q^2 , los que describen la predictibilidad del modelo.

Kutner et. al., definen PRESSp (Suma de error cuadrado predecible) como la medida, por el uso de los valores ajustados para un modelo, de cuan bien se pueden predecir los valores observados y_i . Cada valor ajustado de \hat{y}_i es obtenido al eliminar el i-avo valor del grupo de datos, calculando la función de regresión para el subgrupo remanente con i-1 valores y después usando la función de regresión ajustada para obtener el valor de $\hat{y}_{i(i)}$, predicho para el i-caso. La notación $\hat{y}_{i(i)}$ significa el valor ajustado predicho para el i-caso, y el subíndice (i) que el i-caso fue

omitido cuando se ajustó o calculó la función de regresión. Luego PRESSp se define como:

$$PRESSp = \sum_{i=1}^n (y_i - \hat{y}_{i(i)})^2 \quad (2.21)$$

Considerándose aceptables aquellos modelos que fueron obtenidos con valores de PRESSp bajos.

Todeschini *et al*, 2002, identifican al Coeficiente de Raíz Cuadrada Predictiva (Predictive Squared Correlation Coefficient) o Q^2 , como el estadígrafo con valores menores o iguales a la unidad. Juzga al modelo como aceptable cuando sus valores son de 0.5 o mayores y Q^2 puede tomar aún valores negativos, hasta -1 para casos de muy baja predictibilidad y está definido como:

$$Q^2 = 1 - \frac{\sum_{i=1}^n (y_i - y_{i \text{ calc } LOO})^2}{\sum_{i=1}^n (y_i - \langle y \rangle)^2} \quad (2.22)$$

Donde:

y_i : Valor observado.

$y_{i \text{ calc } LOO}$: Valor calculado por el procedimiento leave-one-out (LOO).

$\langle y \rangle$: Media de los valores $y_{i \text{ calc } LOO}$

La Organización para la Cooperación y el Desarrollo Económico, OCDE, en el documento Guía para la validación de los modelos Relación Cuantitativa Estructura Actividad QSAR (2007), también propone como un parámetro predictivo a la desviación estándar del error predictivo, SDEP, la que está definida por:

$$SDEP = [\sum_{i=1}^n (y_i - \hat{y}_i)^2 / n]^{1/2} = [PRESSp / n]^{1/2} \quad (2.23)$$

n: Número de datos u observaciones.

Para los modelos de regresión lineal múltiple, Montgomery, Peck y Vining, 2012, señalan a los siguientes entre los principales supuestos:

1. El término error tiene una varianza constante. Cuando los términos del error tienen varianza constante u homogénea se denomina propiedad de homocedasticidad y en este caso el modelo propuesto explica bien a la variable dependiente. Los residuos son graficados en las ordenadas, los valores predichos en las abscisas y el gráfico debe presentar una nube de puntos sin ningún patrón.

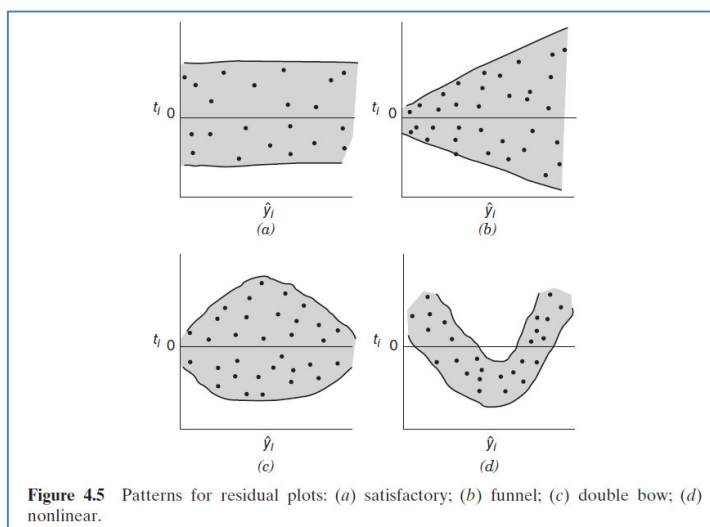


Fig. N° 6. Homocedasticidad y Varianza Residual. (Fuente: Montgomery et al.)

2. Los errores están normalmente distribuidos. Kutner *et al.*, reportan que la evaluación gráfica de la normalidad puede ser realizada al graficar los residuos contra los valores esperados normalizados. Esto mostrará a los puntos alineados en una recta cuando se cumple la normalidad y gráficamente es equivalente a un plot Q –Q.

3. Los errores no están correlacionados. El e-Handbook of Statistical Methods (2013) señala que los gráficos de los residuos frente al valor de la respuesta ajustada deben producir una nube de puntos dispersas alrededor del cero.

2.3.11 Reporte QSPR

La salida típica de los datos de los programas QSPR incluye ecuaciones de correlación que involucran a coeficientes de correlación (R o R^2), tests de significado estadístico, descriptores adecuados y sus correspondientes coeficientes de regresión. Para seleccionar la mejor correlación generalmente se aplica la validación del modelo ya sea mediante técnicas de validación cruzada o por valores de respuesta predictiva para el grupo de datos.

2.3.12 Validación del modelo

La validación es un aspecto importante del modelo relación cuantitativa estructura-actividad (QSAR). Cousins y Mackay, 2000 y Tienpont, 2004, afirman que esta validación se realiza para examinar la estabilidad interna y predicción de los modelos QSAR, los cuales son extensibles en forma válida a los modelos denominados QSPR en la que una propiedad fisicoquímica es investigada.

Kutner, et. al., 2005, refieren que la validación es una parte útil y necesaria del proceso de construcción de modelos y que la validez del modelo alude a la estabilidad y razonabilidad de los coeficientes de regresión; la certeza y uso de la función de regresión y; a la posibilidad de generalizar inferencias extraídas del análisis de regresión.

A su vez Montgomery, et. al., 2012, señalan que se debe distinguir entre la verificación de la adecuación del modelo y la validación del modelo. Ellos indican que la verificación de la adecuación del modelo incluye el análisis residual, las pruebas de falta de ajuste, la búsqueda de observaciones de gran influencia y otros análisis internos que investigan el ajuste del modelo de regresión a los datos disponibles; mientras que la validación del modelo permite establecer si el modelo funcionará con éxito en el entorno o ambiente operacional previsto. A partir de ello se puede concluir que el proceso de validación es una etapa necesaria

cuando el modelo será utilizado, exclusivamente, para fines de predicción.

El análisis de las hipótesis de los modelos de regresión lineal múltiple se basa en los supuestos de aleatoriedad de los errores, definiéndolos como el valor observado menos el valor calculado. Si esto se cumple, ninguna información explicativa debe encontrarse en el error o, gráficamente, no se deberá esperar tendencias en los gráficos de los errores. La principal suposición es que el modelo propuesto explique el comportamiento total de la variable observada y que sólo lo aleatorio, similar a cualquier fenómeno del mundo real, permanece como un sobrante en el valor del error. Si se observa capacidad explicativa o de predicción en los errores, se debe entender que los predictores están perdiendo o han perdido algo de información predictiva.

Tanto Montgomery et al. (2012), así como Kutner et al. (2005), coinciden en señalar que los criterios o técnicas principales para la validación de los modelos de regresión lineal son tres:

1. Colección de nuevos datos para examinar el modelo y su capacidad predictiva
2. Comparación de resultados con expectativas teóricas, resultados empíricos anteriores y los obtenidos por simulación
3. Uso de una muestra retenida para verificar el modelo y su capacidad predictiva

Al comparar los valores obtenidos, empleando el modelo, con los valores observados de datos que no fueron tomados en cuenta en la construcción del modelo, es la forma más difundida para la validación del modelo. Sin embargo, la valoración de la técnica de validación depende de la disponibilidad de datos, así como del uso final del modelo.

A diferencia de Kutner et. al., 2005 que señalan que desde el punto de vista exclusivamente teórico es poco lo que se puede obtener para la

validación de los modelos, Montgomery et. al., 2012, indican que, para este fin, se pueden emplear las métricas VIF y $R^2_{\text{predictora}}$.

Montgomery et. al., 2012, precisan que VIF, denominado factor de inflación de la varianza, mide el efecto combinado de la dependencia entre la regresión y la varianza de ese término donde uno o más valores de VIF grandes indican múltiple colinealidad, además si el valor de VIF es superior a 10 denota múltiple colinealidad.

Se define VIF como:

$$VIF_i = r^{ii} = 1 / (1 - R^2_i), \quad i = 1, \dots, p \quad (2.24)$$

Donde:

R^2_i : Coeficiente de correlación múltiple de regresión entre los x_i y los ($p-1$) predictores.

Murray Leigh, et. al., 2012, citando a Belsey et. al., 1980 y a O'Brien, 2007, mencionan que no hay un criterio claro y definido para considerar como valores de VIF altos o bajos y que se debe tratar con mucho cuidado la presencia de valores de VIF de 10 o 30, para retirar uno o más predictores y que se deben considerar otros factores; como por ejemplo, el tamaño de la muestra. Del mismo modo Murray et. al., 2012, afirman que, dependiendo de la naturaleza de la investigación, se puede obviar el criterio de VIF.

Montgomery et. al., 2012, definen al coeficiente de correlación cuadrado predictora $R^2_{\text{predictora}}$ como la métrica que explicará la variabilidad de la variable independiente dentro de un nuevo grupo de datos. Los valores de esta métrica se encontrarán entre 0 y 1, considerándose como la mejor cuando el valor se encuentre más cerca de la unidad.

$$R^2_{\text{predictora}} = 1 - \frac{\sum e_i^2}{SS_T} \quad (2.25)$$

Donde:

$\sum e_i^2$: Suma de los cuadrados de errores del grupo de predicción.

SS_T : Suma corregida de los cuadrados de las respuestas en el grupo de predicción.

La OCDE, op. cit., señala las condiciones que los modelos QSAR deben considerar para la validación:

1. Un punto final definido. Este se refiere a las condiciones en que el modelo se aplica o usa. Ejemplo: presión, temperatura, etc.
2. Un algoritmo no ambiguo. Se debe emplear un algoritmo unívoco definido.
3. Un dominio de aplicabilidad definido. El campo sobre el que modelo trabaja o se aplica debe estar explícitamente indicado.
4. Medidas apropiadas de la bondad, fortaleza y capacidad de predicción del modelo. El análisis del modelo será mediante ANOVA.
5. La explicación del modelo mediante una interpretación mecanística.

CAPITULO 3: METODOLOGIA

3.1 Recopilación de Ésteres Ftálicos y Selección de sus Valores de Presión de Vapor

La selección de los valores de presión de vapor de los ésteres ftálicos, objeto del presente estudio, fue realizada a partir de la investigación hecha por Staples *et al.*, 1997. Estos junto a sus abreviaturas se muestran en la Tabla N°3 y debido a que, originalmente, sus valores fueron expresados

Tabla N°3. Abreviatura y Nombre de los Ésteres Ftálicos.

ABREVIATURA	ESTER FTALICO
DMP	Dimetil Ftalato
DEP	Dietil Ftalato
DAP	Dialil Ftalato
DPP	Dipropil Ftalato
DnBP	Di-n-Butil Ftalato
DIBP	Diisobutil Ftalato
BBP	Butilbencil Ftalato
DHP	Dihexil Ftalato
DnOP	Di-n-Octil Ftalato
BOP	Butil 2-Etilhexil
610P	Di(n-Hexil, n-Octil, n-Decil)Ftalato
DEHP	Di (2-Etilhexil) Ftalato
DIOP	Diisooctil Ftalato
DINP	Diisononil Ftalato
DIDP	Diisodecil Ftalato
D711P	Di (Heptil, Nonil, Undecil) Ftalato
DUP	Diundecil Ftalato
DTDP	Ditridecil Ftalato

(Fuente: Staples et. al. 1997)

en mmHg se les ha considerado junto a aquellos de las unidades SI; es decir miliPascal (mPa). En la Tabla N°4 se observa que para cada éster ftálico se reporta diversidad de valores de presión de vapor; por lo que los valores de esta propiedad han sido tabulados de manera individual para cada tipo de

éster. En esta Tabla (Nº4) se aprecia la diversidad de valores de presión de vapor asignados para un mismo éster, muchos de los cuales varían en diferentes órdenes de magnitud. En virtud de esta complejidad, Staples et. al. aplicaron el método de juicio experto, a este conjunto de datos (Cousins y Mackay, 2000), el cual es un protocolo derivado de la política de seguridad debido a accidentes ocurridos en plantas nucleares y que combina enfoques matemáticos y conductuales para cuantificar la incertidumbre como una probabilidad subjetiva experta con significado operacional en la que se aplica la inversión probabilística para obtener distribuciones sobre variables no observables y que evalúa la dependencia entre cantidades inciertas (Rasmussen, 1975; Cooke y Gossens, 2000; Cooke, 2012).

Tabla N°4. Valores de Presión de Vapor de Ésteres Ftálicos

PV a 25°C	Di-n-butil Ftalato (DnBP)											
mmHg	9.3x10 ⁻⁵ (35)	1.7x10 ⁻⁵ (32)	1.9x10 ⁻⁵ (29)	2.08x10 ⁻⁵ (30)	2.7x10 ⁻⁵ (27)	3.5x10 ⁻⁵ (34)	3.6x10 ⁻⁵ (36)	4.1x10 ⁻⁵ (30)	7.3x10 ⁻⁵ (2)	1.2x10 ⁻⁴ c	2.5x10 ⁻⁴ d	2.7x10 ⁻⁵ *
mPa	1.24	2.266	2.533	2.773	3.6	4.666	4.8	5.466	9.733	15.999	33.331	3.6

PV a 25°C	Dietil Ftalato (DEP)								
mmHg	4.8x10 ⁻⁵ (27)	3.9x10 ⁻⁴ (28)	6.1x10 ⁻⁴ (28)	8.1x10 ⁻⁴ c	1.2x10 ⁻³ (35)	1.65x10 ⁻³ (2)	5.0x10 ⁻³ d	1.0x10 ⁻³ *	
mPa	6.399	51.996	81.327	107.991	159.987	219.982	666.612	133.322	

PV a 25°C	Dimetil Ftalato (DMP)					
mmHg	1.65x10 ⁻³ (2)	1.8x10 ⁻³ (27)	5.4x10 ⁻³ (38)	5.5x10 ⁻³ c	2.9x10 ⁻² d	2.0x10 ⁻³ *
mPa	219.982	239.98	719.941	733.273	3866.35	266.645

PV a 25°C	Butilencil Ftalato (BBP)					
mmHg	8.7x10 ⁻⁷ (35)	7.7x10 ⁻⁶ d	8.6x10 ⁻⁶ (14)	9.0x10 ⁻⁶ (2)	9.1x10 ⁻⁵ c	5.0x10 ⁻⁶ *
mPa	0.116	1.027	1.147	1.2	12.132	0.667

PV a 25°C	Di-n-Octil Ftalato (DOP)					
mmHg	2.2x10 ⁻⁷ (27)	<3.4x10 ⁻⁷ (35)	3.4x10 ⁻⁷ d	1.3x10 ⁻⁵ c	1.9x10 ⁻⁴ (29)	1.0x10 ⁻⁷ *
mPa	0.029	0.045	0.045	1.733	25.865	0.013

Tabla N°4. Valores de Presión de Vapor de Ésteres Ftálicos (cont)

PV a 25°C	Diisooctil Ftalato (DIOP)					
mmHg	$2.0 \times 10^{-10} c$	$3.4 \times 10^{-7} (35)$	$1.4 \times 10^{-6} d$	$1.4 \times 10^{-6} (29)$	$5.6 \times 10^{-6} (2)$	$1.0 \times 10^{-6} *$
mPa	0	0.045	0.187	0.187	0.747	0.133

PV a 25°C	Diisodecil Ftalato (DIDP)					
mmHg	$5.1 \times 10^{-8} (36)$	$5.6 \times 10^{-8} (35)$	$< 5.0 \times 10^{-7} (2)$	$3.7 \times 10^{-8} d$	$< 1.0 \times 10^{-10} c$	$< 5.0 \times 10^{-7} *$
mPa	0.007	0.007	0.067	0.005	0	0.067

PV a 25°C	Diundecil Ftalato (DUP)				
mmHg	$< 1.0 \times 10^{-10} c$	$1.2 \times 10^{-9} a$	$< 5.0 \times 10^{-7} (2)$	$5.3 \times 10^{-7} (35)$	$< 5.0 \times 10^{-7} *$
mPa	0	0	0.67	0.071	0.067

PV a 25°C	Dihexil Ftalato (DHP)				
mmHg	$1.9 \times 10^{-8} c$	$1.8 \times 10^{-6} (34)$	$1.4 \times 10^{-5} (2)$	$1.2 \times 10^{-5} d$	$5.0 \times 10^{-6} *$
mPa	0.003	0.24	1.867	1.6	0.667

PV a 25°C	Ditridecil Ftalato (DTDP)			
mmHg	$2.5 \times 10^{-11} d$	$< 1.0 \times 10^{-10} c$	$< 5.0 \times 10^{-7} (2)$	$< 5.0 \times 10^{-7} *$
mPa	0	0	0.067	0.067

PV a 25°C	Di(Heptil, Nonil, Undecil) Ftalato (711P)			
mmHg	$5.0 \times 10^{-8} d$	$2.8 \times 10^{-7} (35)$	$< 5.0 \times 10^{-7} (2)$	$< 5.0 \times 10^{-7} *$
mPa	0.007	0.037	0.067	0.067

PV a 25°C	Diisononil Ftalato (DINP)			
mmHg	$< 1.0 \times 10^{-10} c$	$2.3 \times 10^{-7} d$	$5.4 \times 10^{-7} (2)$	$< 5.0 \times 10^{-7} *$
mPa	0	0.031	0.072	0.067

PV a 25°C	Butil 2-Hexil Ftalato (BOP)		
mmHg	$1.1 \times 10^{-7} (35) *$	$2.4 \times 10^{-5} d$	$1.2 \times 10^{-4} c$
mPa	0.015	3.2	15.999

PV a 25°C	Di(n-Hexil, n-Octil, n-Decil) Ftalato (610P)		
mmHg	$3.4 \times 10^{-7} d$	$4.9 \times 10^{-4} c$	$4.9 \times 10^{-6} (2) *$
mPa	0.045	65.328	0.653

PV a 25°C	Dialil Ftalato (DAP)		
mmHg	$3.7 \times 10^{-5} c$	$1.16 \times 10^{-3} d$	$1.6 \times 10^{-4} (35) *$
mPa	4.933	154.654	21.332

Tabla N°4. Valores de Presión de Vapor de Ésteres Ftálicos (cont)

PV a 25°C	Di (2-Etilhexil) Ftalato (DEHP)												
mmHg	4.1x10 ⁸ (33)	4.5x10 ⁸ (12)	7.2x10 ⁸ (35)	9.8x10 ⁸ (11)	2.8x10 ⁷ (29)	3.3x10 ⁷ (36)	<3.4x10 ⁷ (34)	3.8x10 ⁷ (27)	7.1x10 ⁷ (32)	6.4x10 ⁶ (2)	1.4x10 ⁶ d	1.4x10 ⁴ c	1.0x10 ⁷ e
mPa	0.005	0.006	0.010	0.013	0.036	0.044	0.045	0.051	0.095	0.853	0.187	18.665	0.013

PV a 25°C	Diisobutil Ftalato (DIBP)	
mmHg	1.8x10 ⁻⁶ c	5.8x10 ⁻⁴ d
mPa	0.24	77.327

PV a 25°C	Dipropil Ftalato (DPP)	
mmHg	8.9x10 ⁻⁵ c	1.04x10 ⁻³ d
mPa	11.866	138.655

(Fuente: datos tomados de Staples et. al., 1997)

Los resultados de este procedimiento aplicado por Staples et. al., se muestran en la Tabla N°5. En ella sólo se han considerado aquellas especies con valores definidos de presión de vapor y de los 18 ésteres, inicialmente considerados, 2 no tienen valores recomendados y de los 16 restantes 5 no tienen valores definidos; es decir sus valores son menores de 5×10^{-7} mmHg. Si se restan estos 7 con valores no definidos quedan 11 ésteres con valores definidos y recomendados, donde sus números CAS y SMILES se indican en la Tabla N°6 y sus estructuras son expuestas en la Figura N°7.

Tabla N°5. Presiones de Vapor, Masa molar y fórmula global de Ésteres Ftálicos

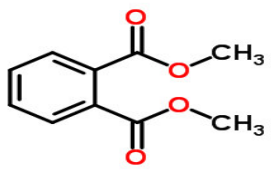
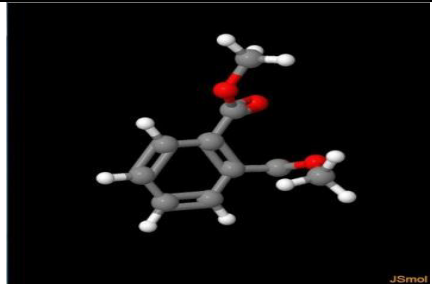
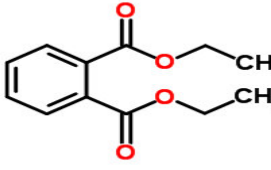
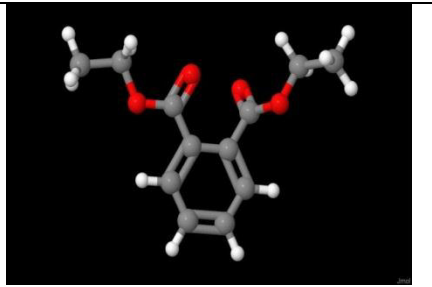
Ester Ftálico	PV (mmHg, ~25 °C)	PV (mPa, ~25 °C)	Masa Molar	Fórmula Global
DMP	2.0x10 ⁻³ *	266.645	194.186	C ₁₀ H ₁₀ O ₄
DEP	1.0x10 ⁻³ *	133.322	222.240	C ₁₂ H ₁₄ O ₄
DAP	1.6x10 ⁻⁴ (35)*	21.332	246.262	C ₁₄ H ₁₄ O ₄
DBP	2.7x10 ⁻⁵ *	3.600	278.348	C ₁₆ H ₂₂ O ₄
BBP	5.0x10 ⁻⁶ *	0.667	312.365	C ₁₉ H ₂₀ O ₄
DHP	5.0x10 ⁻⁶ *	0.667	334.456	C ₂₀ H ₃₀ O ₄
DOP	1.0x10 ⁻⁷ *	0.013	390.564	C ₂₄ H ₃₈ O ₄
610P	4.9x10 ⁻⁶ (2)*	0.653	390.564	C ₂₄ H ₃₈ O ₄
DIOP	1.0x10 ⁻⁶ *	0.133	390.564	C ₂₄ H ₃₈ O ₄
DEHP	1.0x10 ⁻⁷ *	0.013	390.564	C ₂₄ H ₃₈ O ₄
BOP	1.1x10 ⁻⁷ (35)*	0.015	334.456	C ₂₀ H ₃₀ O ₄

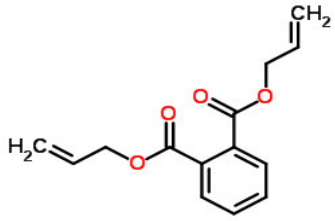
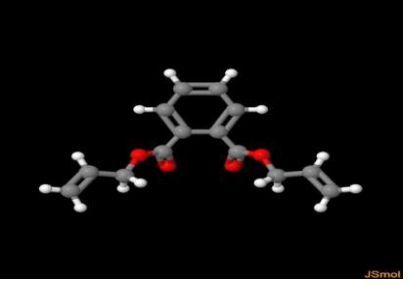
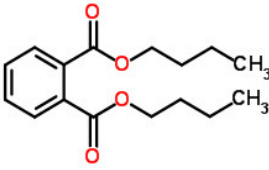
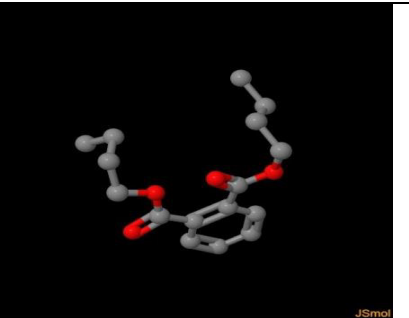
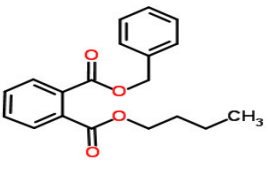
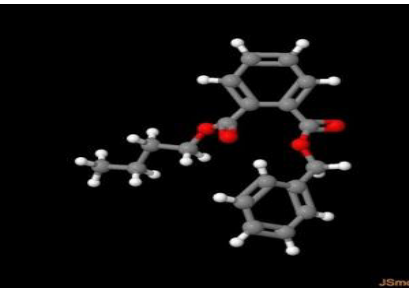
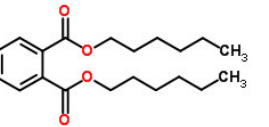
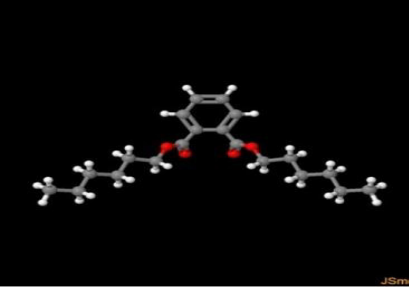
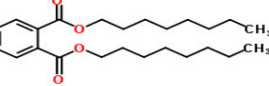
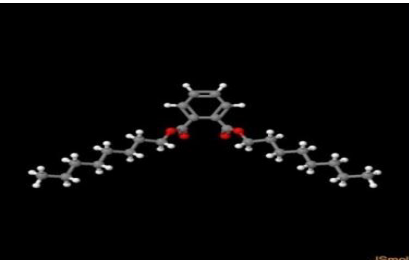
(Fuente: modificada a partir de datos tomados de Staples et. al. [Ref])

Tabla N°6. Identificación de los CAS y SMILES de los ftalatos

SIGLA	ESTER FTALICO	CAS	SMILES
DMP	Dimetil ftalato	131-11-3	<chem>COC(=O)C1=CC=CC=C1C(=O)OC</chem>
DEP	Dietyl ftalato	84-66-2	<chem>CCOC(=O)C1=CC=CC=C1C(=O)OCC</chem>
DAP	Dialil ftalato	131-17-9	<chem>C=CCOC(=O)C1=CC=CC=C1C(=O)OCC=C</chem>
DBP	Di-n-butil ftalato	84-74-2	<chem>CCCCOC(=O)C1=CC=CC=C1C(=O)OCCCC</chem>
BBP	Butilbencil ftalato	85-68-7	<chem>CCCCOC(=O)C1=CC=CC=C1C(=O)OCC2=CC=CC=C2</chem>
DHP	Dihexil ftalato	84-75-3	<chem>CCCCCCOC(=O)C1=CC=CC=C1C(=O)OCCCCCC</chem>
DOP	Di-n-octil ftalato	117-84-0	<chem>CCCCCCCCOC(=O)C1=CC=CC=C1C(=O)OCCCCCCCC</chem>
610P	Di(n-hexil, n-octil, n-decil) ftalato	25724-58-7	<chem>CCCCCCCCCOC(=O)C1=CC=CC=C1C(=O)OCCCCCCC</chem>
DIOP	Diisooctil ftalato	27554-26-3	<chem>CC(C)CCCCOC(=O)C1=CC=CC=C1C(=O)OCCCCC(C)C</chem>
DEHP	Di(2-etilhexil) ftalato	117-81-7	<chem>CCCC(CC)COC(=O)C1=CC=CC=C1C(=O)OCC(CC)CCCC</chem>
BOP	Butil 2-etilhexil ftalato	85-69-8	<chem>CCCC(CC)COC(=O)C1=CC=CC=C1C(=O)OCCCC</chem>

(Fuente: propia)

Ester	Formula desarrollada	Estructura
DMP		
DEP		

DAP	 <chem>C=CCOC(=O)c1ccccc1C(=O)OCC=C</chem>	 JSmol
DBP	 <chem>CCCCOC(=O)c1ccccc1C(=O)OCCCC</chem>	 JSmol
BBP	 <chem>CCCCOC(=O)c1ccccc1C(=O)OCC2=CC=CC=C2</chem>	 JSmol
DHP	 <chem>CCCCCCOC(=O)c1ccccc1C(=O)OCCCCCC</chem>	 JSmol
DOP	 <chem>CCCCCOC(=O)c1ccccc1C(=O)OCCCCC</chem>	 JSmol

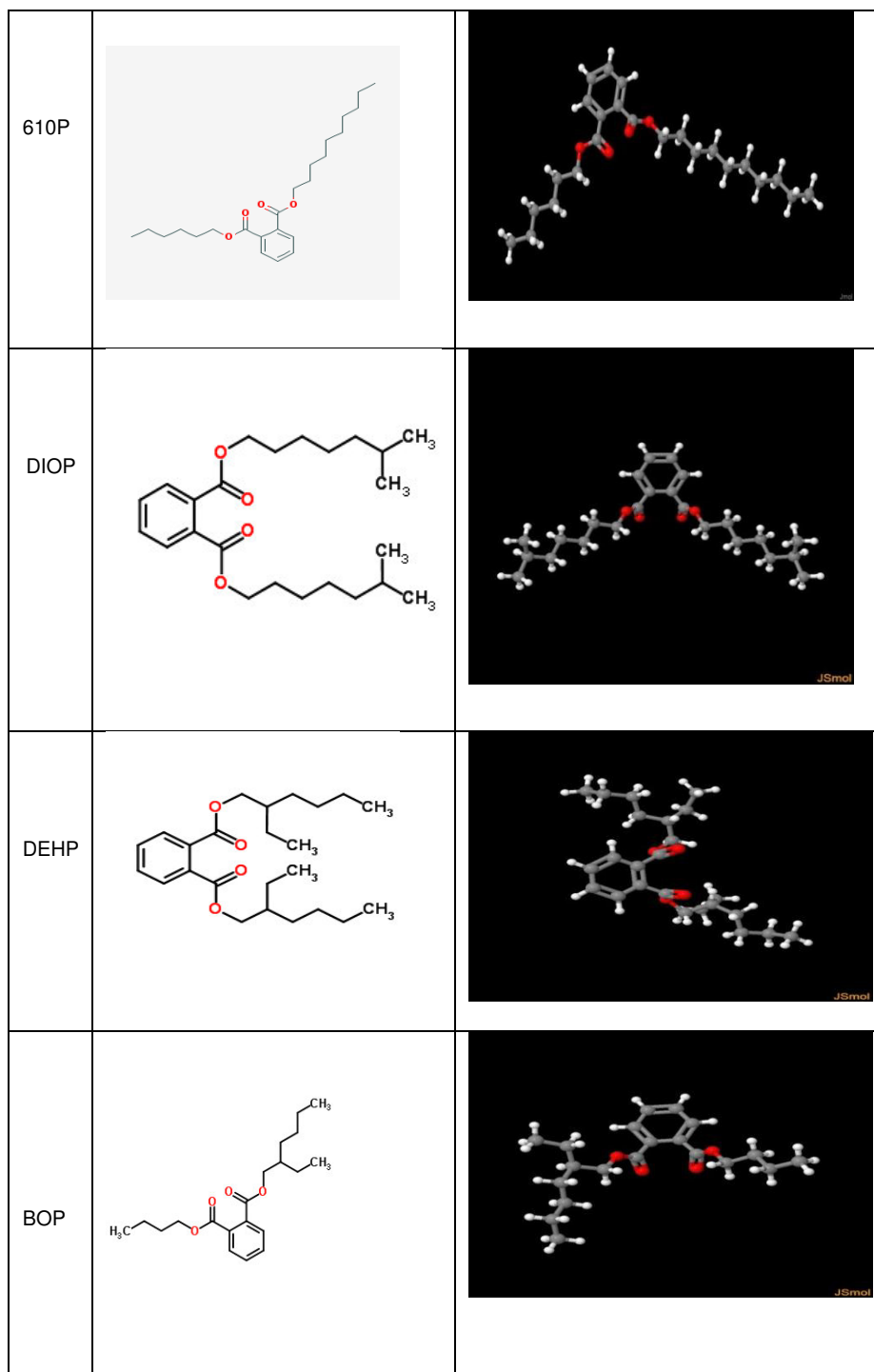


Fig.N°7 . Fórmulas desarrolladas y estructura de Ftalatos. (Fuente: jmol)

3.2 Cálculo de los Valores de los Descriptores Moleculares

Se ingresan en el software E-Dragon 1.0 las representaciones SMILES correspondientes a los ftalatos seleccionados descritos en la Tabla N° 6.

Para cada tipo de descriptor, el software E-Dragon 1.0 ejecuta el algoritmo específico de cada uno de ellos y los resultados los presenta en formato de texto de salida imprimible (Fig. N°8). Para el descriptor RDFu, el algoritmo que ejecuta E-Dragon 1.0 está dado por la ecuación 2.1 y para los demás descriptores RDF por la ecuación 2.3. Los valores de los descriptores se presentan en la Tabla N° 8. En los Anexo 2 se muestra el ejemplo de cálculo de obtención de los descriptores moleculares RDFu, RDFm, RDFv, RDFe y RDFp para el ftalatos DEHP.

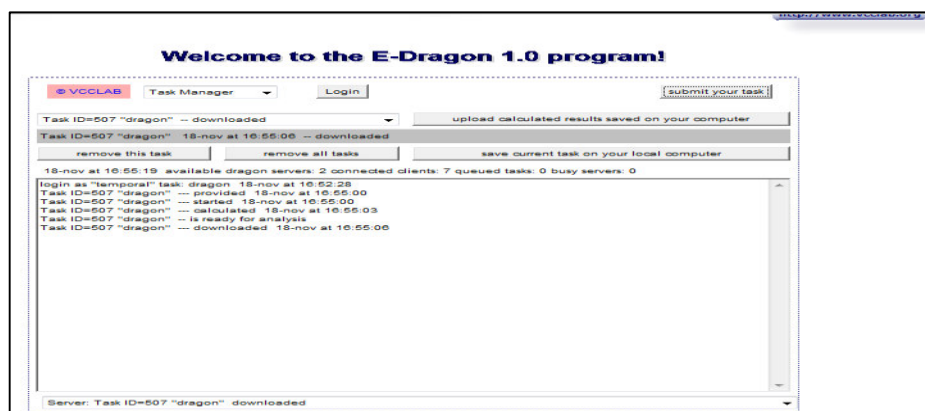


Fig.N°8. Formato de texto de salida imprimible del software E-Dragon 1.0.
(Fuente: E-Dragon 1.0)

Tabla N°7

. Descriptores RDFu proporcionados por E-Dragon 1.0.Grupo 1

No.	MOL_ID	RDF010u	RDF015u	RDF020u	RDF025u	RDF030u	RDF035u	RDF040u	RDF045u	RDF050u	RDF055u	RDF060u	RDF065u	RDF070u	RDF075u	RDF080u	RDF085u
1	C20H30O4	13.696	14.648	7.133	63.724	28.578	7.081	16.171	11.813	49.502	2.567	15.257	26.689	8.549	35.933	8.471	

(Fuente: E-Dragon 1.0)

Tabla N°8. Valores descriptores RDF calculados por E-Dragon 1.0

RDFu

Table with 20 columns (FALATO, RDF00u, RDF01u, ..., RDF15u) and 10 rows (DMP, DEP, DAP, BPP, BBP, DHP, DOP, DEHP, BOP). Values range from 0 to 15.92.

RDFm

Table with 20 columns (FALATO, RDF00m, RDF01m, ..., RDF15m) and 10 rows (DMP, DEP, DAP, BPP, BBP, DHP, DOP, DEHP, BOP). Values range from 0.00 to 12.21.

RDFv

Table with 20 columns (FALATO, RDF00v, RDF01v, ..., RDF15v) and 10 rows (DMP, DEP, DAP, BPP, BBP, DHP, DOP, DEHP, BOP). Values range from 0.00 to 12.54.

RDFe

Table with 20 columns (FALATO, RDF00e, RDF01e, ..., RDF15e) and 10 rows (DMP, DEP, DAP, BPP, BBP, DHP, DOP, DEHP, BOP). Values range from 0.00 to 16.25.

RDFp

Table with 20 columns (FALATO, RDF00p, RDF01p, ..., RDF15p) and 10 rows (DMP, DEP, DAP, BPP, BBP, DHP, DOP, DEHP, BOP). Values range from 0.00 to 15.92.

3.3 Determinación del Modelo

Los valores de los descriptores obtenidos mediante la aplicación de E-Dragon 1.0 son ingresados al software BuildQSAR; como se muestra en las Tablas N° 9 y 10, los cuales son procesados y sus resultados, en formato de salida se muestran en la Tabla N° 11. El formato de salida exhibe sus resultados presentando a los conjuntos de descriptores asociados con estadísticos diversos escogiendo entre a ellos a los que tengan mayor valor de F y menor s en un proceso de selección visual (screening). Realizada la selección de los conjuntos de descriptores con mayores valores de F y menores de s se ingresan estos descriptores al software BuildQSAR seleccionando la opción de presentar una ecuación lineal multivariable. Los resultados para cada conjunto de descriptores son presentados en el formato que es mostrado en las Tablas N° 12 y 13 de las cuales se escoge aquel cuyo resultado que aproxima mejor las presiones de vapor calculadas con las observadas, incluyendo en este paso de verificación a aquellos falatos que no se consideraron para la obtención de valores de descriptores moleculares.

El modelo regresional elegido es aquel donde los valores de presión de vapor, proporcionados por BuildQSAR, son los más cercanos a los experimentales.

Tabla N° 9. BuildQSAR. DMP, DAP, DBP, BBP

ID	Structure	Comment	Y1	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20	X21	X22	X23	X24	X25	X26	X27	X28	X29	X30
1	DMP		266.6448	4.798	5.504	4.085	15.008	3.23	5.582	5.624	8.001	12.99	2.092	8.355	5.583	7.985	2.878	1.38	0.859	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	DAP		21.331584	7.071	7.547	7.668	20.128	7.274	9.379	10.233	6.856	18.971	3.595	15.262	11.931	6.812	8.56	7.202	9.437	3.154	5.061	3.909	1.743	3.268	0.959	1.663	0.544	0.988	0	0	0	0	0
3	DBP		3.5997048	10.137	10.99	5.605	43.951	17.434	7.081	13.371	11.07	32.164	2.39	12.474	15.397	10.019	20.642	6.479	15.898	11.329	8.555	13.077	11.11	4.851	4.535	7.668	5.999	2.851	1.285	2.593	1.999	0	0
4	BBP		0.666612	9.659	10.733	6.173	38.98	10.343	8.697	12.209	15.062	27.14	8.05	13.274	14.648	14.666	12.614	15.228	17.004	12.418	10.494	12.052	11.947	4.549	3.717	6.609	3.656	4.172	2.14	3.04	3.271	2.186	0.001

	X30	X31	X32	X33	X34	X35	X36	X37	X38	X39	X40	X41	X42	X43	X44	X45	X46	X47	X48	X49	X50	X51	X52	X53	X54	X55	X56	X57	X58	X59	X60	X61	X62	X63	X64	X65	X66	X67		
1	0	0.429	6.094	0.448	11.413	3.733	3.772	2.775	1.716	6.165	0.227	3.05	2.535	0.501	0.223	0.01	0.006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0.62	8.138	0.696	14.327	3.876	6.594	3.79	1.71	10.028	0.397	6.074	5.328	1.461	4.558	0.897	3.519	1.019	0.352	0.835	0.843	0.398	0.019	0.066	0.004	0.007	0	0	0	0	0	0	2.118	6.677	1.93	11.859	2.373	3.44	3.086	
3	0	0.878	11.578	0.546	17.35	3.834	3.947	4.169	1.988	13.473	0.299	4.705	7.786	1.409	7.097	1.784	6.021	2.249	1.054	2.73	0.769	0.171	1.874	1.189	0.238	0.02	0.908	0.215	0.014	0	0	3.034	10.116	1.376	16.076	2.969	2.393	4.153		
4	0.001	0.837	11.322	0.596	19.057	3.803	4.11	4.109	5.923	13.204	2.475	4.802	6.303	3.09	4.014	6.578	6.264	5.421	3.742	2.207	0.644	1.136	1.324	1.917	0.737	0.678	0.102	0.355	0.111	0.019	0	2.891	9.86	1.554	17.408	2.348	2.87	3.375		

Tabla N° 10. BuildQSAR. DOP, 610P, DIOP, BOP, DEHP (cont)

	X99	X100	X101	X102	X103	X104	X105	X106	X107	X108	X109	X110	X111	X112	X113	X114	X115	X116	X117	X118	X119	X120	X121	X122	X123	X124	X125	X126	X127	X128	X129	X130	X131	X132
1	66.808	2.919	16.285	44.44	8.706	58.586	8.368	22.553	26.18	12.002	39.392	12.79	18.183	23.763	24.347	22.117	14.83	16.358	16.782	19.783	12.31	11.729	6.57	17.326	2.858	29.958	6.86	2.571	7.393	3.129	26.179	0.695	5.847	18.338
2	66.845	2.918	16.27	44.444	8.712	58.595	8.374	20.593	29.732	12.928	42.742	11.932	20.523	20.204	20.604	23.835	14.363	13.725	15.491	17.758	12.807	11.336	6.57	17.326	2.858	29.957	6.858	2.571	7.394	3.126	26.193	0.695	5.837	18.332
3	59.666	8.144	22.585	34.942	10.552	45.895	9.568	23.69	20.265	14.399	35.331	16.013	26.413	21.219	29.845	22.008	20.417	19.791	16.002	17.877	12.693	11.442	6.57	17.327	2.871	31.52	8.76	5.411	8.45	5.106	23.571	2.547	7.202	14.43
4	48.698	10.416	22.435	30.947	17.704	32.267	13.384	26.244	18.81	12.824	19.763	13.92	13.621	9	11.424	11.176	7.471	2.822	6.035	4.861	2.271	0.629	5.215	13.671	2.413	24.327	7.353	4.22	8.147	5.255	18.308	2.556	10.006	12.675
5	55.028	18.944	31.976	47.915	28.562	31.817	22.691	39.723	29.459	29.228	29.617	24.465	32.47	21.949	17.963	20.6	15.506	8.238	11.266	9.613	4.839	3.785	6.57	17.327	2.956	31.868	11.088	7.071	9.602	8.35	21.574	5.573	12.092	20.36

	X16	X117	X118	X119	X120	X121	X122	X123	X124	X125	X126	X127	X128	X129	X130	X131	X132	X133	X134	X135	X136	X137	X138	X139	X140	X141	X142	X143	X144	X145	X146	X147	X148	X149	X150
1	358	16.782	19.783	12.31	11.729	6.57	17.326	2.858	29.958	6.86	2.571	7.393	3.129	26.179	0.695	5.847	18.338	2.902	23.232	2.884	8.902	8.711	4.244	14.206	4.806	6.772	9.228	8.712	8.126	4.367	7.063	6.114	7.232	4.618	3.471
2	725	15.491	17.758	12.807	11.335	6.57	17.326	2.858	29.957	6.858	2.571	7.394	3.126	26.193	0.695	5.837	18.332	2.909	23.239	2.884	8.483	9.774	4.583	17.017	4.458	7.231	8.243	7.753	6.32	4.134	5.601	6.087	7.545	4.548	3.324
3	791	16.002	17.877	12.693	11.442	6.57	17.327	2.871	31.52	8.76	5.411	8.45	5.106	23.571	2.547	7.202	14.43	3.466	18.597	3.084	9.059	7.98	4.665	15.443	5.912	9.673	10.166	10.717	8.193	6.009	8.054	5.043	6.016	4.616	3.351
4	22	6.035	4.861	2.271	0.629	5.215	13.671	2.413	24.327	7.353	4.22	8.147	5.255	18.308	2.556	10.006	12.675	6.056	13.831	4.092	11.153	7.391	4.932	6.978	4.233	6.091	4.337	4.153	3.464	2.083	1.244	2.015	1.502	0.558	0.259
5	38	11.256	9.613	4.839	3.785	6.57	17.327	2.956	31.868	11.088	7.071	9.602	8.35	21.574	5.573	12.092	20.36	11.209	14.044	7.848	17.606	11.891	10.186	10.796	9.134	13.275	9.742	6.103	6.978	5.154	2.864	3.482	2.772	1.532	1.671

Tabla N°11. Coeficientes asociados a cada descriptor, a través de sus parámetros estadísticos R, F, s y Q² calculados por BuildQSAR.

	X-1	X-2	X-3	R	s	F	Q ²	SPress	SDep
1	X3	X7	X144	0.998	0.034	90.008	0.381	0.437	0.218
2	X3	X10	X139	1.000	0.003	9128.082	0.997	0.030	0.015
3	X3	X13	X26	0.997	0.043	54.629	0.643	0.332	0.166
4	X3	X15	X23	0.995	0.053	35.667	0.439	0.416	0.208
5	X3	X15	X113	1.000	0.016	408.701	0.859	0.209	0.104
6	X3	X16	X53	0.999	0.025	164.358	0.210	0.493	0.247
7	X3	X26	X103	0.994	0.060	28.370	0.391	0.433	0.217
8	X3	X35	X139	1.000	0.000	12891134...	1.000	0.001	0.000
9	X3	X38	X84	1.000	0.015	436.192	0.944	0.132	0.066
10	X3	X38	X144	0.996	0.051	38.835	0.489	0.397	0.198
11	X3	X40	X86	1.000	0.005	4048.098	0.993	0.046	0.023
12	X3	X40	X145	1.000	0.017	355.629	0.873	0.197	0.099
13	X3	X53	X75	0.999	0.028	129.451	0.635	0.335	0.168
14	X3	X53	X76	1.000	0.015	459.988	0.804	0.246	0.123
15	X3	X53	X135	0.999	0.026	153.106	0.117	0.522	0.261
16	X3	X53	X136	0.999	0.021	241.959	0.551	0.372	0.186
17	X3	X73	X83	1.000	0.007	2422.060	0.984	0.070	0.035
18	X3	X97	X144	0.999	0.020	252.632	0.834	0.226	0.113
19	X3	X100	X139	1.000	0.006	3176.440	0.991	0.054	0.027
20	X3	X105	X113	0.999	0.024	178.124	0.843	0.220	0.110
21	X3	X105	X143	1.000	0.007	2194.584	0.984	0.071	0.036
22	X3	X113	X133	1.000	0.017	352.114	0.914	0.163	0.081
23	X3	X133	X143	1.000	0.003	13836.163	0.997	0.031	0.015
24	X5	X47	X54	1.000	0.008	1454.208	0.950	0.124	0.062
25	X5	X56	X78	0.999	0.018	303.390	0.402	0.429	0.215
26	X5	X59	X108	1.000	0.014	529.134	0.853	0.213	0.106
27	X5	X88	X93	1.000	0.016	380.691	0.665	0.321	0.161
28	X5	X88	X125	0.999	0.029	120.033	0.003	0.554	0.277
29	X5	X108	X119	1.000	0.007	1857.777	0.954	0.120	0.060
30	X6	X28	X145	1.000	0.011	832.096	0.972	0.094	0.047
31	X6	X34	X87	0.999	0.018	309.623	0.812	0.241	0.120

(Fuente: BuildQSAR)

Tabla N°12. Modelo de regresión lineal de variable múltiple.

MLR results

QSAR Model | Fitting analysis | Dataset | Graphics | Full analysis

QSAR Equation
=====

$$Y1 = - 50.5872 (\pm 0.2646) X82 - 33.2622 (\pm 0.0973) X101 + 578.9780 (\pm 1.3108)$$

(n = 4 ; R = 1.000 ; s = 0.039 ; F = 16236720.073 ; p = 0.0002 ; Q2 = 1.000 ; SPress = 1.539 ; SDEP = 0.888)

Dataset
=====

No.	Compounds	Y1	X82	X101
1	DMP	266.645	0.000	9.390
2	DAP	21.332	0.118	16.586
3	DBP	3.600	2.273	13.842
4	BBP	0.667	1.654	14.870

*(Fuente: BuildQSAR)***Tabla N°13.** Modelo de regresión lineal de variable múltiple.

MLR results

QSAR Model | Fitting analysis | Dataset | Graphics | Full analysis

QSAR Equation
=====

$$Y1 = - 0.1142 (\pm 0.0196) X12 + 0.1166 (\pm 0.0200) X18 + 0.5491 (\pm 0.0784) X58 + 1.6345 (\pm 0.3296)$$

(n = 5 ; R = 1.000 ; s = 0.005 ; F = 3723.348 ; p = 0.0120 ; Q2 = 0.903 ; SPress = 0.173 ; SDEP = 0.086)

Dataset
=====

No.	Compounds	Y1	X12	X18	X58
1	DNOP	0.013	45.324	11.981	3.934
2	610P	0.653	45.326	12.922	4.901
3	DIOP	0.133	35.265	14.160	1.589
4	BOP	0.015	30.915	13.628	0.594
5	DEHP	0.013	48.156	30.296	0.635

(Fuente: BuildQSAR)

CAPITULO 4: RESULTADOS Y DISCUSIÓN

La evaluación de los valores de los descriptores moleculares generados por E-Dragon 1.0 de las moléculas de los ésteres ftálicos usando BuilQSAR establece que son los descriptores moleculares RDF los que describen mejor el comportamiento de la presión de vapor, proveyendo la mejor aproximación a la presión de vapor observada.

Los mejores resultados se obtuvieron agrupando a los ésteres ftálicos en dos conjuntos y con el menor número de descriptores.

4.1. Conjunto de Ésteres DMP, DAP, DBP, BBP y DEP

Para los ésteres ftálicos DMP, DAP, DBP, BBP y DEP, el modelo de regresión lineal múltiple esta mostrado en la Tabla N° 12 y se expresa mediante la ecuación:

$$P_v = - 50.5872[\text{RDF115v}] - 33.2622[\text{RDF060e}] + 578.9780 \quad (4.1)$$

Los valores correspondientes para los descriptores RDF115v y RDF060e de este grupo de ésteres son mostrados en la Tabla N°14.

Tabla N°14. Descriptores RDF115v, RDF060e

ESTER	PRESION	RDF115v	RDF060e
DMP	266.645	0	9.39
DAP	21.332	0.118	16.586
DBP	3.600	2.273	13.842
BBP	0.667	1.654	14.87
DEP	133.322	0	13.538

(Fuente: E-Dragon 1.0)

Los espectros de los diferentes valores que toma $g(r)$ en la ecuación (4.1) correspondientes a los descriptores RDFv y RDFe para distancias comprendidas entre 1 y 15.5 Å se muestran en las figuras N°9 y N°10. Ver anexos 3 y 4

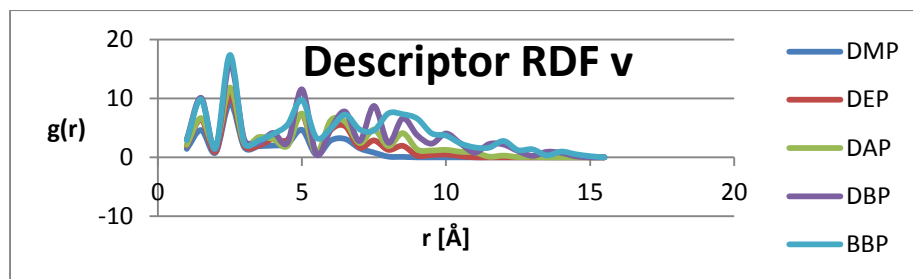


Fig. N°9. Espectros descriptor RDFv de los ésteres DMP-DEP-DAP-DBP-BBP. (Fuente: propia)

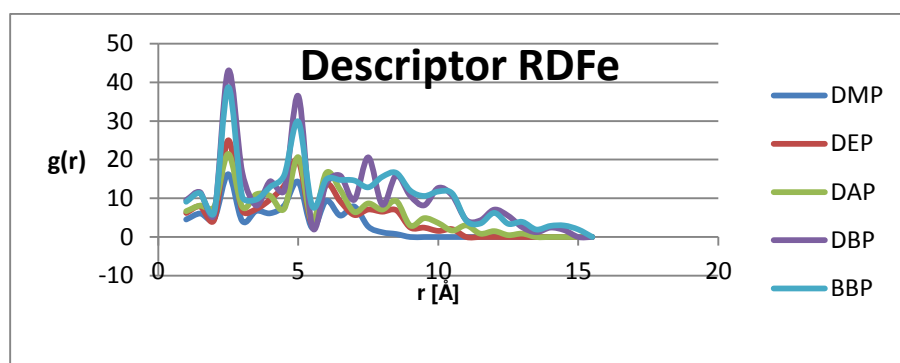


Fig. N°10. Espectros descriptor RDFe de los ésteres DMP-DEP-DAP-DBP-BBP. (Fuente: propia)

Los datos proporcionados en la Tabla N°14 permiten obtener el análisis de varianza mostrado en la Tabla N°15.

Tabla N°15. Análisis de varianza 1

Fuente	GL	Suma de cuadrados	Cuadrados medios	F	Pr > F
Modelo	2	50216.375	25108.187	16269835.435	0.000
Error	1	0.002	0.002		
Total corregido	3	50216.376			

Calculado contra el modelo $Y = \text{Media}(Y)$

(Fuente: propia. Microsoft Excel 2010)

El análisis de la varianza (ANOVA) proporciona los valores de F 16269835 y de significancia $0.00 < 0.05$. Estos valores permiten afirmar que existe dependencia lineal, con un grado de confiabilidad del 95%, entre la presión de vapor y los descriptores moleculares RDF115v y RDF060e.

Tabla N°16. Coeficiente de correlación 1

	R ²	R ² adj
Valor	1	1

(Fuente: propia)

Cuando R² toma el valor de 1 indica que, el modelo como tal explica el 100% de varianza de la presión de vapor. Así mismo, el valor de 1 para R²adj indica que el valor de R² no se verá afectado por la variación del número de variables explicativas o que el número de variables que forman el modelo es el adecuado.

Tabla N°17. Parámetros del modelo 1

Fuente	Valor	Error estándar	t	Pr > t	Límite inferior (95%)	Límite superior (95%)
Intersección	578.978	0.103	5617.973	0.000	577.668	580.287
RDF115v	-50.587	0.021	-2431.641	0.000	-50.852	-50.323
RDF060e	-33.262	0.008	-4347.907	0.000	-33.359	-33.165

(Fuente: propia. Microsoft Excel 2010)

Del análisis de los parámetros del modelo mostrados en la Tabla N°17, el valor del intercepto o intersección correspondiente a la ecuación 4.1 puede interpretarse como el promedio de la presión de vapor correspondiente a las estructuras en base a las que se desarrollaron el modelo. La métrica estadística t nos permite identificar, en el modelo de regresión que el descriptor RDF060e tiene mayor significancia entre las variables explicativas, no obstante, el valor de Pr > |t| indica que todas las variables poseen significancia en el modelo.

Los valores correspondientes a PRESS_p, SDEP y Q²_{LOO}, mostrados en la Tabla N°18, fueron obtenidos mediante el procedimiento de validación cruzada LOO (leave-one-out) y ellos permiten afirmar que el modelo posee capacidad predictiva.

Tabla N°18. Capacidad predictora 1

	PRESSp	SDEP	Q ² _{LOO}
Valor	2.365	0.7689	1

(Fuente: propia)

Aplicación del Modelo obtenido (ecuación 4.1) para el cálculo de la presión de vapor de DEP:

$$P_{vDEP} = - 50.5872[\text{RDF115v}] - 33.2622[\text{RDF060e}] + 578.9780$$

$$P_{vDEP} = - 50.5872[0] - 33.2622[13.538] + 578.9780$$

$$P_{vDEP} = 128.674$$

En la Tabla N°19 se muestran los correspondientes residuales calculados; es decir, valor observado menos valor calculado.

Tabla N°19. RESIDUALES 1 para los DMP, DAP, DBP y BBP

	Obs	Calc	Residual	StDev. Res	Valor esperado normalizado	√StDev. Res
DMP	266.645	266.646	-0.001	-0.025	-0.012	0.159
DAP	21.332	21.322	0.010	0.254	0.012	0.504
DBP	3.6	3.578	0.022	0.560	0.041	0.748
BBP	0.667	0.698	-0.031	-0.788	-0.041	0.888

(Fuente: propia)

En las figuras N° 11, 12 y 13 se muestran los gráficos construidos con los valores de la Tabla N°19, considerando la varianza constante (homocedasticidad), normalidad e independencia de residuos.

En la figura N°10 se muestra una franja de valores o variación constante sin presentar tendencia de los residuos estandarizados graficados frente a los valores calculados. En la figura N°11 los valores de los residuos graficados en función de los valores esperados normalizados se alinean alrededor de una recta y la figura N°12 muestra independencia respecto de los residuos.

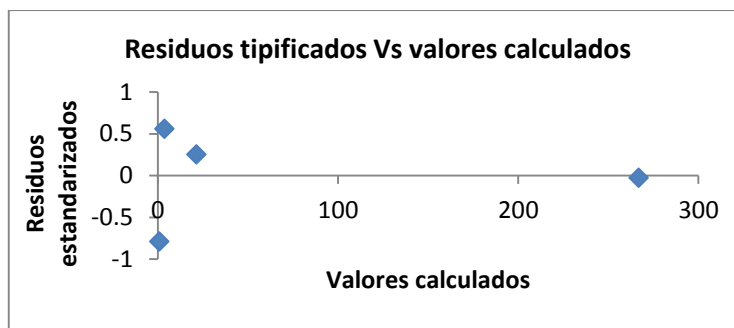


Fig. N°11. Homocedasticidad residuales 1. (Fuente: propia)

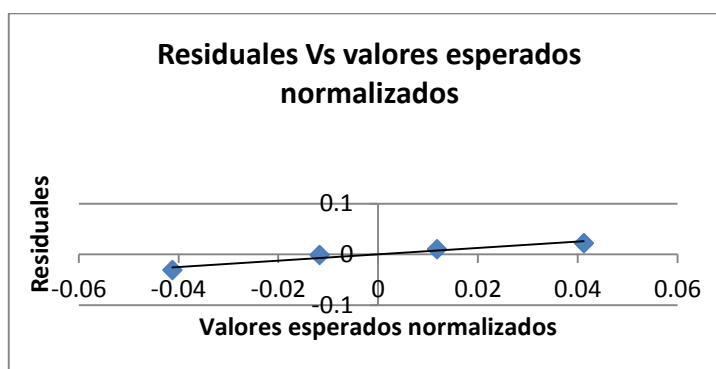


Fig.N°12. Normalidad. (Fuente: propia)

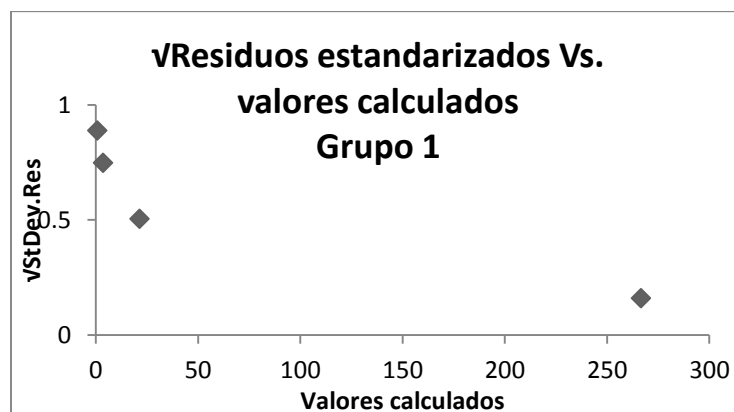


Fig. N°13. Independencia. (Fuente: propia)

En la Tabla N°20 se comparan los valores observados y calculados, obtenidos al aplicar el modelo descrito en la ecuación (4.1).

Tabla N°20. Presión de vapor de los ésteres DMP-DAP-DBP-BBP-DEP

Especie	Observado	Calculado
DMP	266.645	266.646
DAP	21.332	21.322
DBP	3.600	3.578
BBP	0.667	0.698
*DEP	133.322	128.674

(Fuente: propia)

En este cuadro el éster DEP no formó parte de los datos para la construcción del modelo. El valor calculado de 128.674 mPa obtenido por aplicación del modelo representa una diferencia de 4.69 mPa con un error del 3.49% respecto al valor observado.

En la Figura N°14 se grafican los valores observados frente a los calculados por la aplicación del modelo de la ecuación 4.1

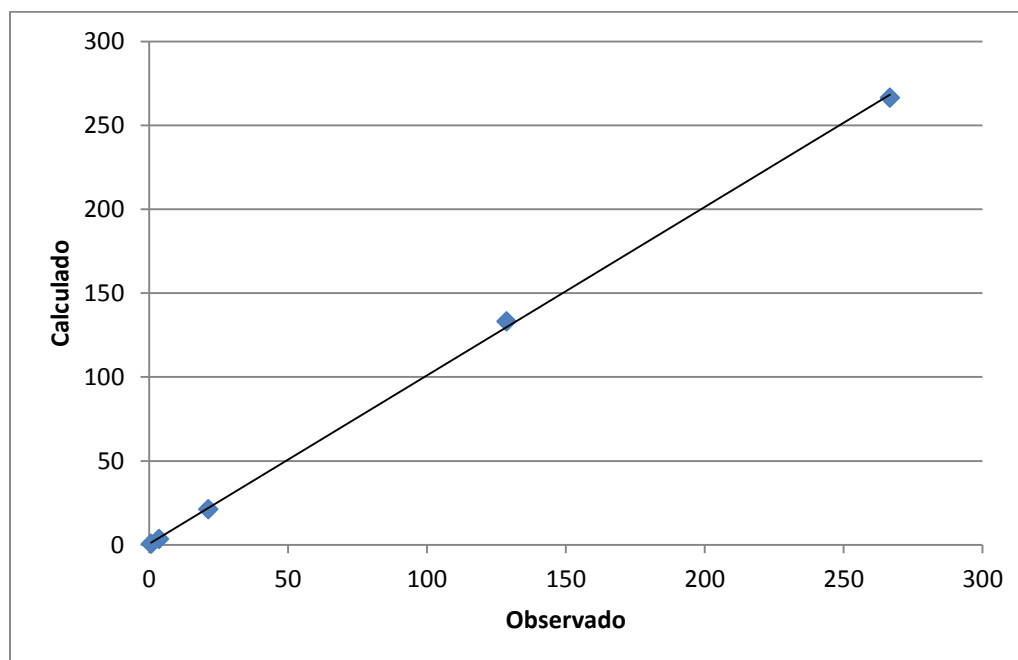


Fig.N°14. Valor Observado Vs. Calculado de los **DMP-DEP-DAP-DBP-BBP**
(Fuente: propia)

4.2 Conjunto de Ésteres DNOP, 610P, DIOP, DEHP, BOP y DHP

De manera similar a lo expresado en 4.1, en el caso de los ésteres ftálicos DNOP, 610P, DIOP, DEHP, BOP y DHP el modelo de regresión lineal múltiple esta mostrado en la Tabla N° 21 y se expresa mediante la siguiente ecuación:

$$P_v = -0.1142[\text{RDF065u}] + 0.1166[\text{RDF095u}] + 0.5491[\text{RDF145m}] + 1.6345 \quad (4.2)$$

Los valores correspondientes para los descriptores RDF065u, RDF095u y RDF145m de este grupo de ésteres son mostrados en la Tabla N°21.

Tabla N°21. Descriptores RDFu, RDFm

ESTER	PRESION	RDF065u	RDF095u	RDF145m
DNOP	0.013	45.324	11.981	3.934
610P	0.653	45.326	12.922	4.901
DIOP	0.133	35.265	14.160	1.589
DEHP	0.013	48.156	30.296	0.635
BOP	0.015	30.915	13.628	0.594
DHP	0.653	26.689	13.861	0.724

(Fuente: E-Dragon 1.0)

De acuerdo a la ecuación (4.2), en las figuras N°15 y N°16 se muestran los espectros que pertenecen a los diferentes valores de $g(r)$ correspondientes a los descriptores RDFu y RDFm, para distancias comprendidas entre 1 y 15.5 Å. Ver anexos 5 y 6.

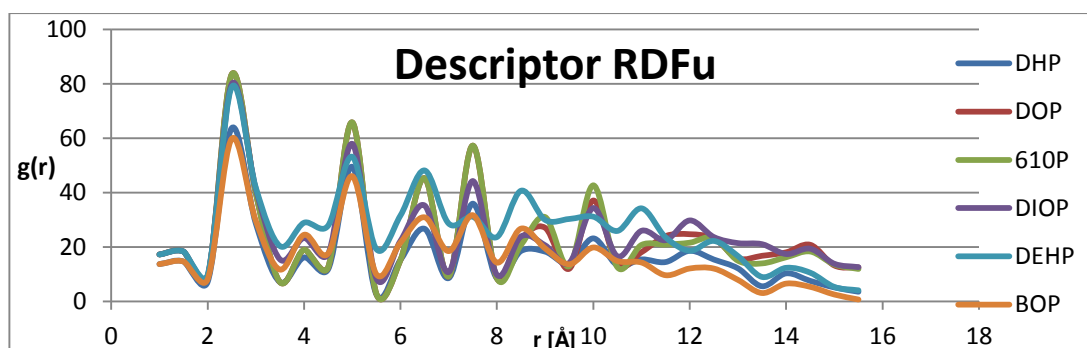


Fig.N°15.Espectro del descriptor RDFu DNOP-610P-DIOP-BOP-DEHP- DHP

(Fuente: propia)

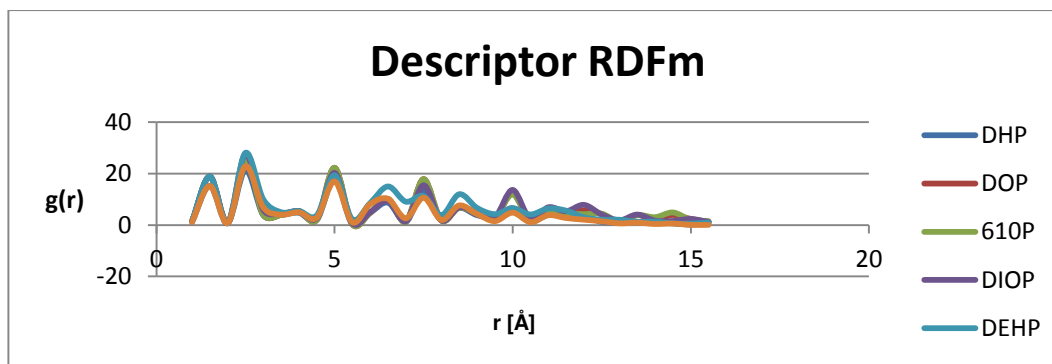


Fig.N°16. Espectro del descriptor **RDFm**. **DNOP-610P-DIOP-BOP-DEHP-DHP**.
(Fuente: propia)

El análisis de varianza mostrado en la Tabla N°22 se obtiene a partir de los datos proporcionados en la Tabla N°21 y en ella se muestra los valores de F, 4324.394, y de significancia $0.01 < 0.05$. Estos permiten afirmar que existe dependencia lineal, con un grado de confiabilidad del 95%, entre la presión de vapor y los descriptores moleculares RDF065u, RDF095u y RDF145m.

Tabla N°22. Análisis de varianza 2

Fuente	GL	Suma de cuadrados	Cuadrados medios	F	Pr > F
Modelo	3	0.308	0.103	4324.394	0.011
Error	1	0.000	0.000		
Total corregido	4	0.308			

Calculado contra el modelo $Y=Media(Y)$

(Fuente: propia, Microsoft Excel 2010)

En la Tabla N°23 vemos que R^2 toma el valor de 0.9999 indicando que el modelo como tal explica el 99.99% de varianza de la presión de vapor. Así mismo, el valor de 99.96 para R^2_{adj} , comparado con el valor de R^2 , indica que el número de variables explicativas o que el número de variables que forman el modelo es el adecuado

Tabla N°23. Coeficiente de correlación 2

	R^2	R^2_{adj}
Valor	0.9999	0.9996

(Fuente: propia)

El análisis de los parámetros del modelo presentado en la ecuación 4.2 se muestran en la Tabla N°24 en la cual el valor de la intersección o intercepto puede interpretarse como el promedio de la presión de vapor, correspondiente a las estructuras en base a las que se desarrollaron el modelo. La métrica estadística t nos permite identificar, en el modelo de regresión que el descriptor RDF145m tiene mayor significancia entre las variables explicativas, no obstante, el valor de $Pr > |t|$ indica que todas las variables poseen significancia en el modelo, como se muestra en la Tabla N°24.

Tabla N°24. Parámetros del modelo 2

Fuente	Valor	Error estándar	t	Pr > t	Límite inferior (95%)	Límite superior (95%)
Intercepción	1.636	0.024	67.963	0.009	1.330	1.941
RDF065u	-0.114	0.001	-79.785	0.008	-0.132	-0.096
RDF095u	0.117	0.001	79.713	0.008	0.098	0.135
RDF145m	0.549	0.006	95.897	0.007	0.476	0.622

(Fuente: propia. Microsoft Excel 2010)

Los valores correspondientes a PRESSp, SDEP, Q^2_{LOO} presentados en la Tabla N°25 fueron obtenidos por el procedimiento de validación cruzada LOO (leave-one-out). Los valores calculados permiten afirmar que el modelo posee capacidad predictiva.

Tabla N°25. Capacidad predictora 2

	PRESSp	SDEP	Q^2_{LOO}
Valor	0.026	0.072	0.92

(Fuente: propia)

Aplicación del Modelo obtenido (ecuación 4.2) para el cálculo de la presión de vapor de DHP:

$$P_{VDHP} = - 0.1142[RDF065u] + 0.1166[RDF095u] + 0.5491[RDF145m] + 1.6345$$

$$P_{VDHP} = - 0.1142[26.689] + 0.1166[13.861] + 0.5491[0.724] + 1.6345$$

$$P_{VDHP} = 0.600$$

En la Tabla N°26 se muestran los residuales calculados (valor observado menos valor calculado)

Tabla N°26. Residuales 2: DNOP, 610P, DIOP, DEHP, BOP

Ester	Obs	Calc	Residual	StDev. Res	Valor esperado normalizado	vStDev. Res
DNOP	0.013	0.015	-0.002	-0.189	-0.003	0.435
610P	0.653	0.656	-0.003	-0.189	-0.003	0.435
DIOP	0.133	0.131	0.002	0.756	0.000	0.869
DEHP	0.013	0.019	-0.006	-0.189	-0.003	0.435
BOP	0.015	0.016	-0.001	-0.567	-0.006	0.753

(Fuente. Propia)

En las figuras 17, 18 y 19 se muestran los gráficos obtenidos en base a los valores de la Tabla N° 26; para la varianza constante (homocedasticidad), normalidad e independencia de residuos.

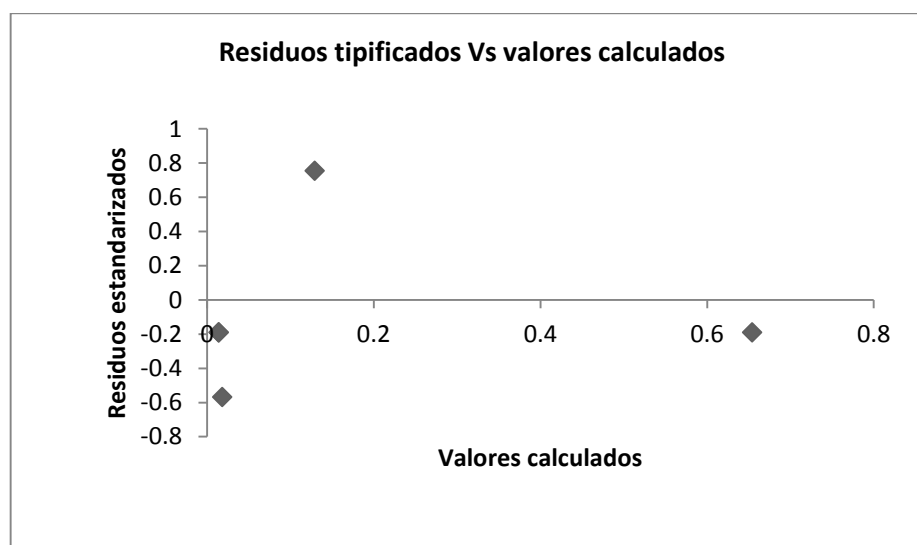


Fig.N°17. Homocedasticidad. Residuales 2 (Fuente: propia)

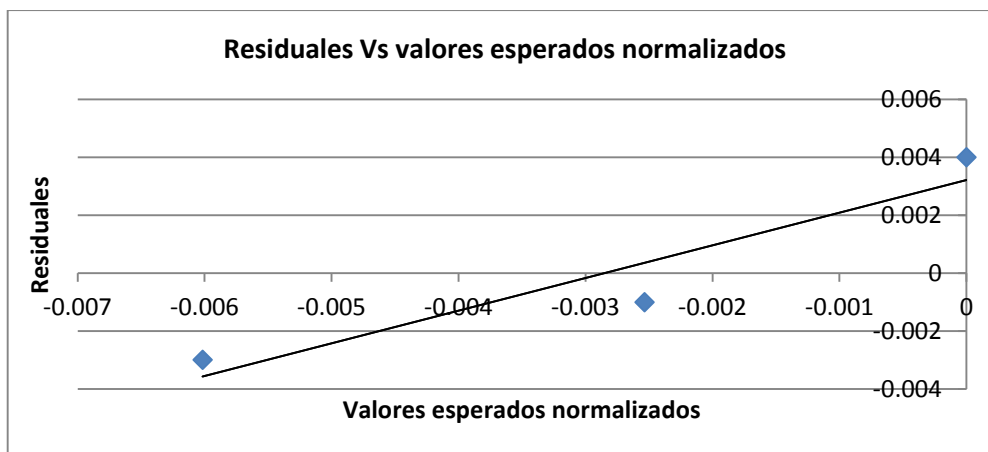


Fig.N°18. Normalidad. Residuales 2. (Fuente: propia)

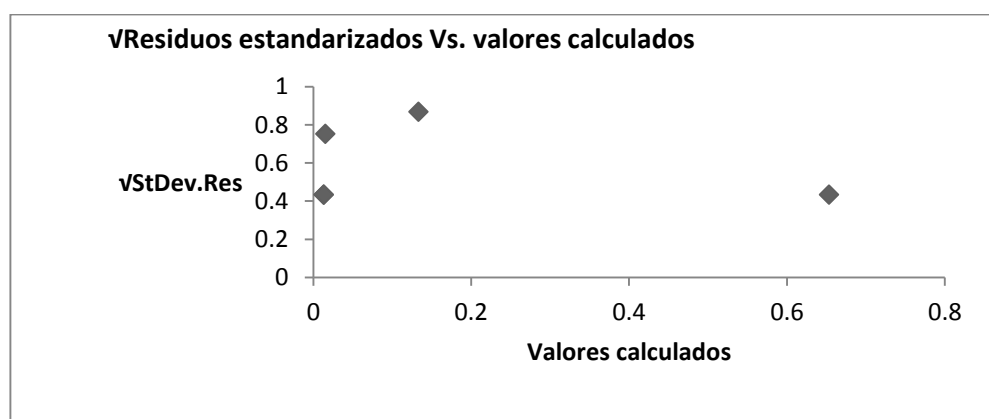


Fig.N°19. Independencia. Residuales 2. (Fuente: propia)

En la figura N°17 son graficados los residuos estandarizados frente a los valores calculados mostrando una franja o variación constante sin presentar tendencia. La figura N°18 muestra los valores de los residuos graficados en función de los normalizados esperados y alineados alrededor de una recta. La figura N°19 muestra la independencia de los residuos.

Aplicando el modelo descrito en la ecuación 4.2 fueron calculados los valores de las presiones de vapor de los ésteres y para su comparación, en la Tabla N°27, se muestran los valores observados y calculados. En ella se observa que el éster DHP no forma parte de los datos para la construcción del modelo ya que, al aplicarlo, el valor calculado de 0.600 mPa representa una diferencia de 0.053 mPa con un error del 8.11% respecto al valor observado.

Tabla N°27. Presión de vapor observada Vs. calculada
DNOP, 610P, DIOP, DEHP, BOP

Especie	Observado	Calculado
DNOP	0.013	0.015
610P	0.653	0.656
DIOP	0.133	0.131
DEHP	0.013	0.019
BOP	0.015	0.016
*DHP	0.653	0.600

(Fuente: propia)

La relación entre los valores observados y los calculados, aplicando el modelo de la ecuación (4.2), son mostrados la figura N° 20.

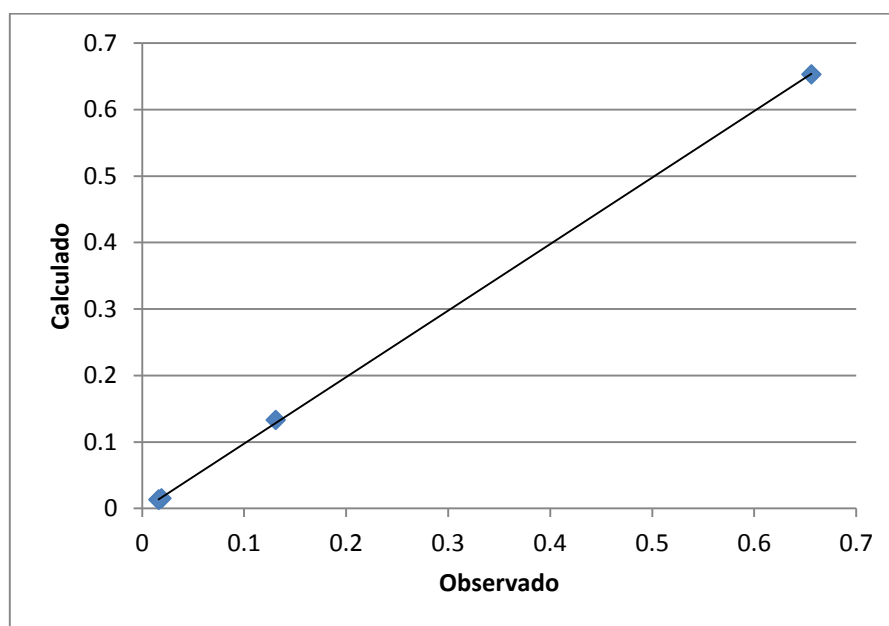


Fig.N°20. Valor observado Vs. calculado de los ésteres. (Fuente: propia)

La Tabla N°28, como resultado global, contiene los valores obtenidos, observados y calculados, de los ésteres DMP, DAP, DBP, BBP y DEP; al aplicar los descriptores moleculares RDF; así como de los ésteres DNOP, 610P, DIOP, DEHP, BOP y DHP al aplicar los descriptores moleculares RDFu y RDFm.

Tabla N°28. Resultado General. Presión de vapor observada Vs calculada.

Ester	Obs	Calc
DMP	266.645	266.646
DAP	21.332	21.322
DBP	3.600	3.578
BBP	0.667	0.698
*DEP	133.322	128.674
DNOP	0.013	0.015
610P	0.653	0.656
DIOP	0.133	0.131
DEHP	0.013	0.019
BOP	0.015	0.016
*DHP	0.653	0.600

(Fuente: propia)

En la figura N°21 se relacionan los valores de las presiones de vapor, observadas y calculadas, correspondientes a los valores mostrados en la Tabla N°28.

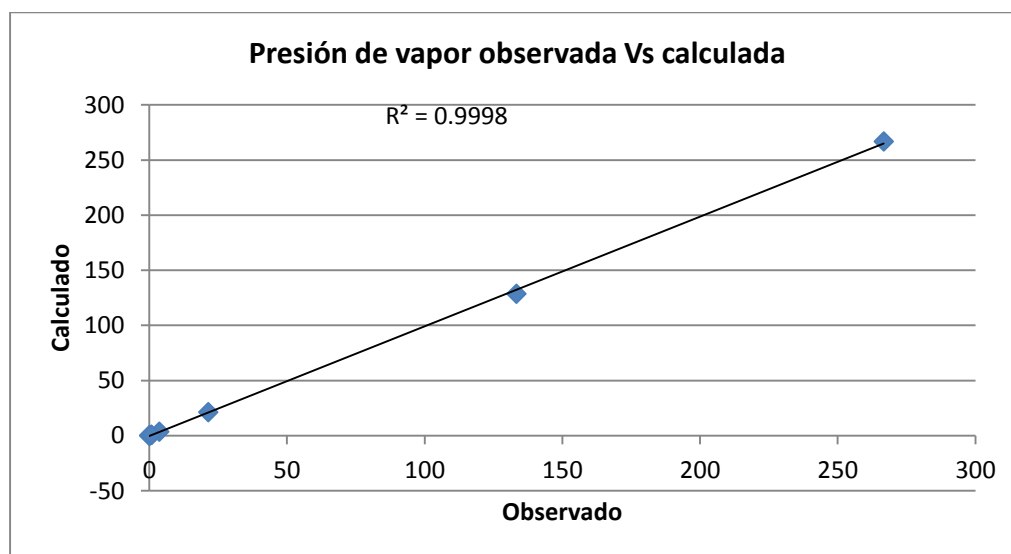


Fig.N°21. Relación de presiones de vapor, obtenidos aplicando diferentes descriptores moleculares. (Fuente: propia)

4.3 Descriptores moleculares determinados

Debe ser considerado que la información que proporcionan los softwares comerciales para descriptores moleculares proveen únicamente el valor numérico del descriptor y; respecto a la divulgación de trabajos sobre la materia, éstos concluyen sus trabajos al publicar la ecuación lineal de

variable múltiple. Sin embargo, es necesario investigar el significado estructural de los resultados. Para ello y considerando que, de acuerdo a la ecuación (2.3), el valor de los descriptores moleculares es la suma de las contribuciones de cada átomo integrante de la molécula a diferentes radios, lo cual requiere expandir o desarrollar el algoritmo de la función de distribución radial (ecuación 2.3) en los radios que indican los descriptores moleculares que conforman el modelo de la ecuación de regresión lineal de variable múltiple; ecuaciones (4.1) y (4.2). En los resultados que se obtienen al desarrollar el algoritmo, mostrados en las Tablas N° 29, 30, 31, 32 y 33, se observa que cada átomo aporta diferente tipo de contribución.

Como criterio de estudio se ordenaron a los átomos de cada éster, de manera jerarquizada y de acuerdo a su contribución porcentual al valor del descriptor. Los resultados se muestran en la Tabla N° 34 para el grupo de ftalatos DAP, DBP, BBP, DMP y DEP y en la Tabla N°35 para el grupo de ésteres DNOP, 610P, DIOP, DEHP, BOP y DHP; en las que se muestra la contribución porcentual de los 5 primeros átomos de cada ftalato. En estas tablas se evidencia la importancia del átomo que se encuentra en primer lugar, a partir del cual es posible presentar un esbozo estructural de acuerdo al radio propuesto por el descriptor, buscando una similitud o patrón estructural en función a la contribución porcentual del átomo que contribuye en mayor proporción al valor del descriptor

En relación al análisis de los parámetros de los modelos obtenidos se debe considerar que el valor del intercepto, de la ecuación del modelo de regresión, se puede atribuir a una correspondencia entre una estructura mínima común entre los ftalatos del grupo y el promedio de la presión de vapor para el grupo. La diferencia de las presiones de vapor específicas para cada éster está dada por la diferencia entre los valores del descriptor para cada molécula.

TablaNº34. Principales átomos contribuyentes. Ftalatos DAP, DBP, BBP, DMP, DEP

	RDF115v			RDF060e		
	Orden	Atomo	%	Orden	Atomo	%
DAP	1	20	58.29	1	4	17.78
(100%-RDFv)	2	18	27.31	2	6	13.79
(63%-RDFe)	3	17	14.40	3	14	13.64
	4	19	0.00	4	12	10.52
	5	1	0.00	5	2	6.83
DBP	1	1	47.41	1	5	22.22
(98%-RDFv)	2	3	44.04	2	15	20.23
(66%-RDFe)	3	20	3.64	3	3	8.19
	4	21	2.50	4	7	7.75
	5	27	0.76	5	14	7.25
BBP	1	1	41.52	1	5	20.16
(95%-RDFv)	2	3	38.82	2	15	18.37
(65%-RDFe)	3	22	8.74	3	16	10.35
	4	32	4.00	4	3	7.43
	5	19	1.80	5	7	7.12
DMP	1	-	-	1	2	31.61
	2	-	-	2	4	11.56
(74%-RDFe)	3	-	-	3	12	11.40
	4	-	-	4	1	10.13
	5	-	-	5	7	9.66
DEP	1	-	-	1	3	21.76
	2	-	-	2	13	20.69
(66%-RDFe)	3	-	-	3	1	8.37
	4	-	-	4	5	7.92
	5	-	-	5	2	6.97

TablaNº35. Principales átomos contribuyentes. Ftalatos DNOP, 610P, DIOP, DEHP, BOP, DHP

	RDF065u			RDF095u			RDF145m		
	Orden	Atomo	%	Orden	Atomo	%	Orden	Atomo	%
DNOP	1	8	10.25	1	5	14.55	1	1	67.96
(29%-RDF065u)	2	7	5.14	2	11	13.70	2	6	12.15
(37%-RDF095u)	3	21	4.45	3	39	8.33	3	4	2.80
(88%-RDF145m)	4	22	4.40	4	23	8.19	4	16	2.46
	5	1	4.40	5	16	7.79	5	7	2.43
610P	1	10	10.27	1	20	14.94	1	3	51.32
(31%-RDF065u)	2	7	6.21	2	13	13.21	2	2	21.57
(54%-RDF095u)	3	9	5.15	3	7	9.75	3	8	9.18
(86%-RDF145m)	4	23	4.46	4	43	8.03	4	6	2.24
	5	5	4.40	5	25	7.89	5	18	1.86
DIOP	1	8	10.25	1	5	14.55	1	6	31.24
(29%-RDF065u)	2	7	5.14	2	11	13.70	2	7	14.74
(53%-RDF095u)	3	21	4.45	3	39	8.33	3	27	10.15
(72%-RDF145m)	4	22	4.40	4	23	8.19	4	22	9.29
	5	1	4.40	5	16	7.79	5	4	6.44
DEHP	1	8	10.77	1	9	9.54	1	2	51.78
(33%-RDF065u)	2	16	7.92	2	8	6.52	2	1	21.57
(35%-RDF095u)	3	19	5.81	3	13	6.27	3	22	11.70
(96%-RDF145m)	4	17	4.37	4	10	6.14	4	23	8.50
	5	1	4.27	5	39	6.07	5	34	2.02
BOP	1	8	15.12	1	3	14.01	1	2	78.82
(39%-RDF065u)	2	19	6.27	2	23	11.71	2	1	8.73
(48%-RDF095u)	3	16	6.20	3	10	7.60	3	22	8.68
(99%-RDF145m)	4	1	5.90	4	7	7.45	4	25	1.34
	5	7	5.68	5	12	7.31	5	28	1.03
DHP	1	6	17.53	1	16	12.68	1	4	60.46
(46%-RDF065u)	2	19	7.62	2	9	12.18	2	2	13.95
(50%-RDF095u)	3	1	7.52	3	3	11.90	3	5	8.65
(92%-RDF145m)	4	24	7.42	4	31	6.81	4	20	5.85
	5	14	6.04	5	8	6.81	5	22	2.63

En la figura N° 22 se muestra la estructura mínima común para los ftalatos DMP, DAP, DBP, BBP y DEP y en la Fig N° 23 la correspondiente a los ftalatos DNOP, 610P, DIOP, DEHP, BOP y DHP.

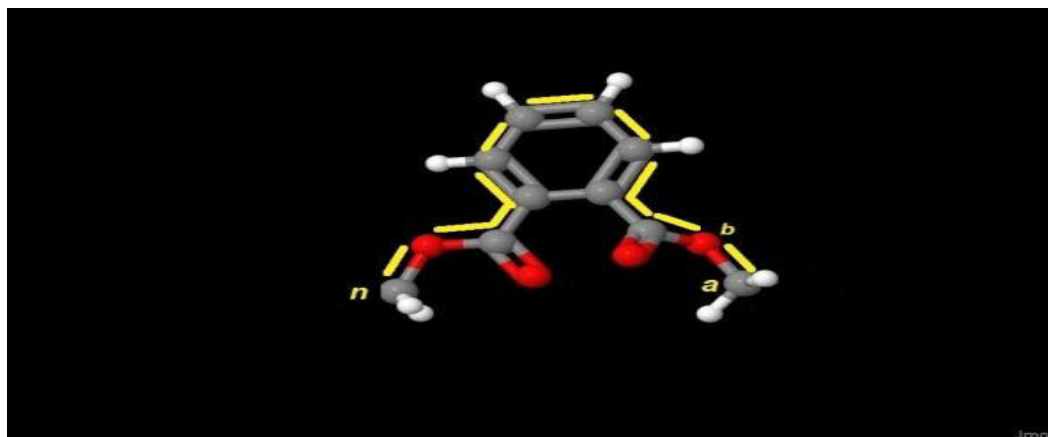


Figura N°22. Estructura mínima común. Ftalatos DMP, DAP, DBP, BBP y DEP (Fuente propia)

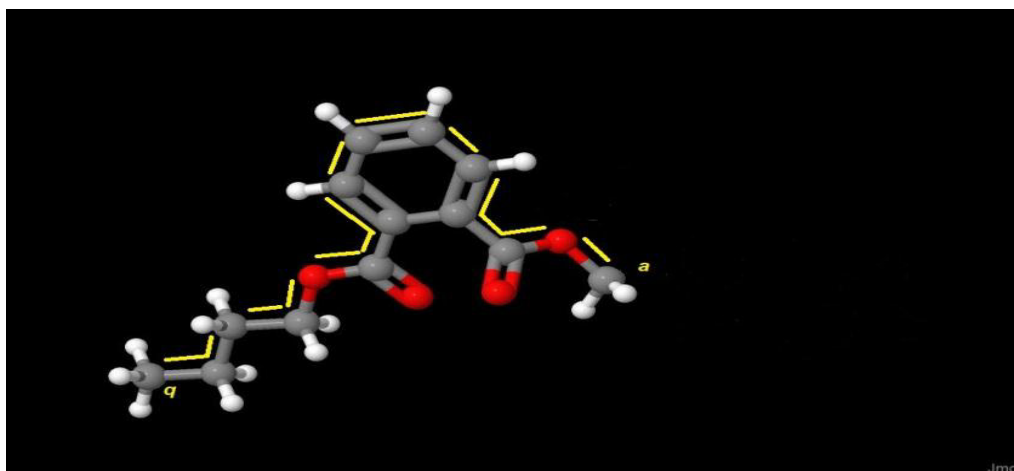


Figura N°23. Estructura mínima común. Ftalatos DNOP, 610P, DIOP, DEHP, BOP y DHP (Fuente propia)

Se aprecia que la estructura mínima común para el grupo 1 de ésteres está dado la posición de los átomos desde **a** hasta **n**, mientras que para el segundo grupo la estructura mínima común está dada por la posición de los átomos de **a** hasta **q**.

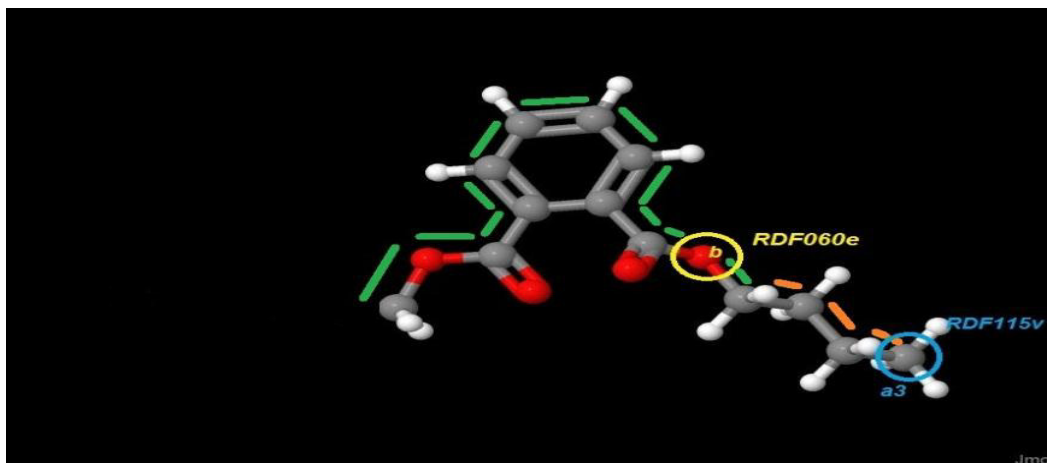


Figura N°24. Modelo contributivo estructural posicional aproximado. Ftalatos DMP, DAP, DBP, BBP y DEP (Fuente. Propia)

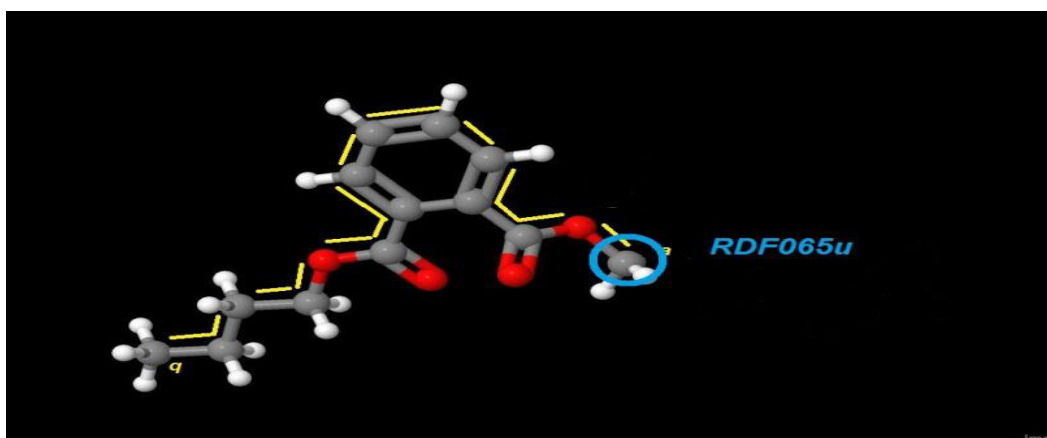


Figura N°25. Modelo contributivo estructural posicional aproximado. Ftalatos DNOP, 610P, DIOP, DEHP, BOP y DHP (Fuente. Propia)

Para el primer grupo se obtiene la Figura N°24 en la cual se aprecia que el valor de la presión de vapor de este grupo de ésteres puede ser atribuida principalmente en primer lugar al átomo en la posición **b** y a sus interacciones ponderadas por electronegatividad con los átomos de la molécula en una vecindad de radio 6.0 Å y también por el átomo en la posición **a3** ponderada por el volumen van der Waals y sus interacciones con la vecindad de átomos en un radio de 11.5 Å .

La presión de vapor disminuye, respecto del promedio dado por la ecuación del modelo, cuando el átomo de oxígeno es ubicado en la posición **b** por el descriptor RDF060e. Con ello, en un radio de 6.0 Å, aparecen interacciones moleculares cuando la presencia de cargas parciales generan dipolos debido a

la diferencia de electronegatividad del átomo de oxígeno respecto a los demás átomos.

La presión de vapor disminuye, respecto del promedio dado por la ecuación del modelo, cuando el átomo de oxígeno es ubicado en la posición **b** por el descriptor RDF060e. Con ello, en un radio de 6.0 Å, aparecen interacciones moleculares cuando la presencia de cargas parciales generan dipolos debido a la diferencia de electronegatividad del átomo de oxígeno respecto a los demás átomos.

En relación al descriptor RDF11.5v, se aprecia que es el átomo de carbono en la posición **a3** es el que más contribuye al valor del descriptor; pero que está ausente en los ftalatos DMP y DEP porque su diámetro es menor que el volumen generado a 11.5 Å, por lo que su presión de vapor, únicamente, esta determinada por el descriptor RDF060e.

En la figura N°23, correspondiente al segundo grupo de ftalatos: DNOP, 610P, DIOP, DEHP, BOP y DHP; debe considerarse que a diferencia del grupo de ftalatos DMP, DAP, DBP, BBP y DEP que no tienen isómeros, está formado por 4 ftalatos que son isómeros de cadena, con la misma composición y masa molar (DNOP, 610P, DIOP y DEHP), y 2 ftalatos con la misma composición y masa molar (BOP y DHP). Es de esperarse, entonces, que al tener una estructura mínima común, generadora de una presión de vapor promedio, se diferencien entre ellos por los valores de los descriptores no ponderados; es decir por las distintas posiciones de los átomos y también por la masa atómica ya que son dos grupos de isómeros con distintas masas molares. Por lo que es posible esperar que los descriptores RDF095u y RDF145m para cada ftalato ubiquen los átomos en distintas posiciones (ver tabla N°31) para obtener diferentes valores de la presión de vapor. En la figura N°25 se muestra la incidencia del descriptor RDF065u, el cual se encuentra comprendido dentro de la estructura mínima común para el grupo de 2 de ftalatos señalando la incidencia del átomo en la posición **a** y en un radio de 6.5 Å, el que contribuye al valor del descriptor RDFu.

De todo lo tratado previamente respecto de la presión de vapor de los ésteres ftálicos, se hace necesario precisar algunas particularidades tales como aquellas que:

- La presión de vapor de los ésteres DMP, DAP, DBP, BBP y DEP fueron calculados al aplicar los descriptores moleculares RDF115v y RDF060e y el modelo de regresión lineal de variable múltiple reporta un valor promedio de 578.978, para su arreglo estructural. La especificación del modelo para la presión de vapor de cada éster, pondera el volumen van der Waals calculado para un radio de 11.5 Å, al aplicar el descriptor RDF115v y aplicando el descriptor RDF060e pondera la electronegatividad de Sanderson calculado para un radio de 6.0 Å; ambos dados por el producto de los coeficientes de regresión de las variables explicativas por los valores específicos de los descriptores de cada éster. El signo negativo de ambos descriptores establece que ambos contribuyen negativamente en el cálculo de la presión de vapor o que la reducen respecto al promedio en una cantidad dada.
- La presión de vapor de los ésteres DNOP, 610P, DIOP, BOP, DEHP y DHP fueron calculados al aplicar los descriptores moleculares RDF065u, RDF095u y RDF145m y el modelo de regresión lineal de variable múltiple reporta un valor promedio de 1.6345, para su arreglo estructural. Los descriptores moleculares no ponderados RDF065u y RDF095u calculados para los radios 6.5 Å y 9.5Å, así como el descriptor RDF145m ponderado por la masa atómica, calculada para un radio de 14.5 Å refinan o especifican el modelo para la presión de vapor de cada éster. El signo positivo de los descriptores RDF 095u y RDF145m indican que, respecto del promedio, contribuyen positivamente al cálculo de la presión de vapor en proporción directa a sus coeficientes de

regresión; mientras que el descriptor RDF065u lo reduce en proporción directa a su coeficiente de regresión.

- Para calcular la presión de vapor de los ésteres ftálicos, expresados a través de las ecuaciones (4.1) y (4.2), los modelos obtenidos en base a descriptores moleculares RDF muestran poseer significancia estadística y que cumplen con las suposiciones hechas para los modelos de regresión lineal de variable múltiple. En los resultados obtenidos por la aplicación de estos modelos se observa una buena correlación entre los valores de la presión de vapor, observados y calculados.
- La aplicación de la ecuación (4.1) en el cálculo de la presión de vapor para el DEP, no considerado en la construcción del modelo expresado por la ecuación (4.1), establece capacidad predictiva del modelo debido a que presenta un porcentaje de error de 3.49% entre el valor de presión de vapor observado y el calculado.
- La aplicación de la ecuación (4.2) en el cálculo de la presión de vapor para el DHP, no considerado en la construcción del modelo expresado por la ecuación (4.2) establece capacidad predictiva del modelo debido a que presenta un porcentaje de error de 8.11% entre el valor de presión de vapor observado y el calculado.
- El modelo de regresión lineal de variable múltiple, expresado en la ecuación (4.2), establece el empleo del descriptor molecular RDFm y de dos descriptores moleculares no ponderados RDFu considerando que los ésteres DOP, 610P, DIOP y DEHP son isómeros de cadena y a que los descriptores RDFu no ponderados, solamente consideran la distribución espacial de los átomos.

CONCLUSIONES

Luego del análisis y los cálculos realizados aplicando descriptores moleculares diferentes, se pudo llegar a las siguientes conclusiones.

- Es posible calcular las presiones de vapor de diferentes ésteres ftálicos al relacionarlos con su estructura molecular. Éstos fueron calculados, únicamente, a partir de sus estructuras químicas y poseen significado físico comparable a los valores experimentales o recomendados.
- Los descriptores moleculares de la Función de Distribución Radial (RDF) calculan adecuadamente las presiones de vapor de los ésteres ftálicos.
- Las combinaciones de los descriptores moleculares representan la interacción estructural de las moléculas, resumida en la ecuación de regresión lineal múltiple.
- La combinación de los descriptores seleccionados dentro del modelo de regresión representan a las interacciones moleculares que caracterizan a la presión de vapor de los ésteres estudiados.
- El cálculo de la presión de vapor de los ésteres ftálicos, aplicando descriptores RDF y examinados por Staples *et al.*, son comparables a los valores denominados recomendados cuando se agrupan en dos blocks.

REFERENCIAS BIBLIOGRAFICAS

- 1.- Abraham Michael H. and Acree William E., (2015), Physicochemical and Biochemical properties for the dialkyl phthalates. *Chemosphere* 119, 871-880.
- 2.- Atkins Peter. Physical Chemistry (2006) Julio de Paula. Eighth Edition. Oxford University Press 2006 p. 118.
- 3.- Basak Subhash C. and Mills Denise, (2005) Development of quantitative structure-activity relationship models for vapor pressure estimation using computed molecular descriptors, *Natural Resources Research Institute*, University of Minnesota Duluth. USA. ARKIVOC 2005 (X), 308-320.
- 4.- Bidleman TF, *Environ Sci Technol*, 1988, 22, 361-367.
- 5.- Cooke R.M., Gossens L.J., (2000) *European Commission Nuclear Science and Technology: Procedures Guide for Structured Expert Judgment*, Delft University of Technology, Netherlands.
- 6.- Cooke, R.M., (2012) Uncertainty Analysis Comes to Integrated Assessment Models for Climate Change and Conversely. *Climatic Change* 117, 467- 479.
- 7.- Cousins I. and Mackay D., (2000) Correlating the physicalchemical properties of phthalate esters using the 'three solubility' approach. *Chemosphere* 41, 1389 -1399.
- 8.- Defoe D., *et al*, (1990) *Env Tox Chem*, 9: 623-36.
- 9.- De Oliveira, Daniel Barbosa; Gaudio, Anderson Coser, 2001. BuildQSAR: A new computer program for QSAR analysis. *Quantitative Structure-Activity Relationships* 19(6): 599-601. March
- 10.- Faulon Jean-Loup, Bender Andreas editors. Handbook of chemoinformatics algorithms /, (Chapman & Hall/CRC mathematical and computational biology series).2010].

- 11.- Gasteiger Johann and Engel Thomas (Eds.) *Chemoinformatics: A Textbook*. Edited by. 2003)
- 12.- Gramatica Paola in collaboration with Dr. Pamela Pilutti. Evaluation of Different Statistical Approaches For The Validation Of Quantitative Structure – Activity Relationships. The European Commission - Joint Research Centre Institute for Health & Consumer Protection - ECVAM 21020 Ispra (VA). Italy.2004.
- 13.- Hammett Louis P.. The Effect of Structure upon the Reactions of Organic Compounds. Benzene Derivatives. *J. Am. Chem. Soc.*, 1937, 59 (1), pp 96–103
- 14.- Hansch Corwin , Fujita Toshio . ρ - σ - π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *J. Am. Chem. Soc.*, 1964, 86 (8), pp 1616–162
- 15.- Hemmer Markus Christof. Radial Distribution Function in Computational Chemistry – Theory and applications. Tesis doctoral Universidad Friedrich-Alexander de Erlangen-Nürnberg.2007
- 16.- Hemmer Markus C., Steinhauer Valentin, Gasteiger Johann. Deriving the 3D structure of organic molecules from their infrared spectra., *Vibrational Spectroscopy* 19 (1999) 151-164
- 17.- Katritzky Alan R., Wang Yilin, Sild Sulev, and Tamm Tarmo, Karelson Mati QSPR Studies on Vapor Pressure, Aqueous Solubility, and the Prediction of Water-Air Partition Coefficients.. *Journal of Chemical Information and Modeling*. pp720-725· July 1998.
- 18.- Katritzky Alan, Lobanov Victor S., Karelson Mati. QSPR: The Correlation and Quantitative Prediction of Chemical and Physical Properties from Structure. *Chem. Soc. Rev.*, 1995, 24, 279-287

- 19.- Katritzky Alan R., Slavov Svetoslav H., Dobchev Dimitar A., Karelson Mati Rapid QSPR model development technique for prediction of vapor pressure of organic compounds. *Computers and Chemical Engineering* 31 (2007), pp. 1123–1130.
- 20.- Kutner M., Nachtsheim J., Neter J., Li W. *Applied Linear Statistical Models*. Fifth Edition. McGraw-Hill Irwin. (2005)
- 21.- Lange's Handbook of Chemistry. Fifteenth Edition. McGraw-Hill, Inc.USA,1999.
- 22.- Levine Ira N. *Química Cuántica*. 5^{ta} Edición. Pearson Educación, S.A., Madrid, 2001, p. 538.
- 23.- Li Li, Jianxin Hu, and Yuh-Shan Ho. Global performance and trend of QSAR/QSPR research: A bibliometric analysis. *Mol. Inf.* 2014, 33, pp. 655 – 668.
- 24.- Liang C., Gallagher D., QSPR Prediction of Vapor Pressure from Solely Theoretically-Derived Descriptors, *J. Chem. Inf. Comput. Sci.*, 1997, pp. 3-8.
- 25.- Liew Chin Yee and Yap Chun Wei “Current Modeling Methods Used in QSAR/QSPR” *Statistical Modelling of Molecular Descriptors in QSAR/QSPR*. First Edition. Published 2012 by Wiley-VCH Verlag GmbH & Co. KGaA
- 26.- Lu Chunhui. Prediction of Environmental Properties in Water-Soil-Air Systems for Phthalates. *Bull. Environ. Contam. Toxicol.* (2009) 83:168-173.
- 27.- Mackay Donald, Wan Ying Shui, Kuo-Ching Ma, Sum Chi Lee. *Handbook of physical-chemical properties and environmental fate for organic chemicals*. 2nd Ed..Published in 2006 by CRC Press Taylor & Francis Group. Boca Raton, FL 3487-2742. USA. 2006

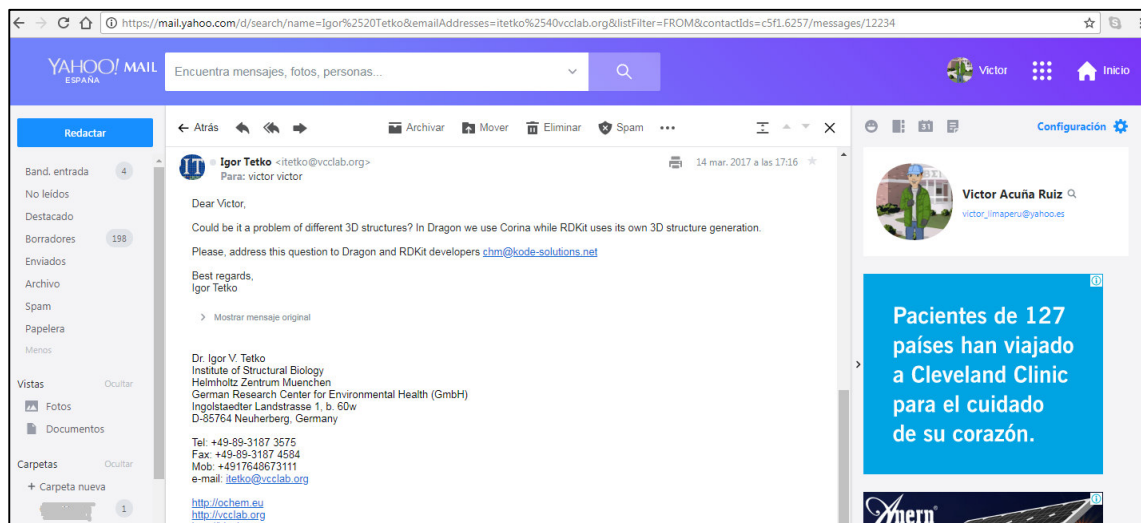
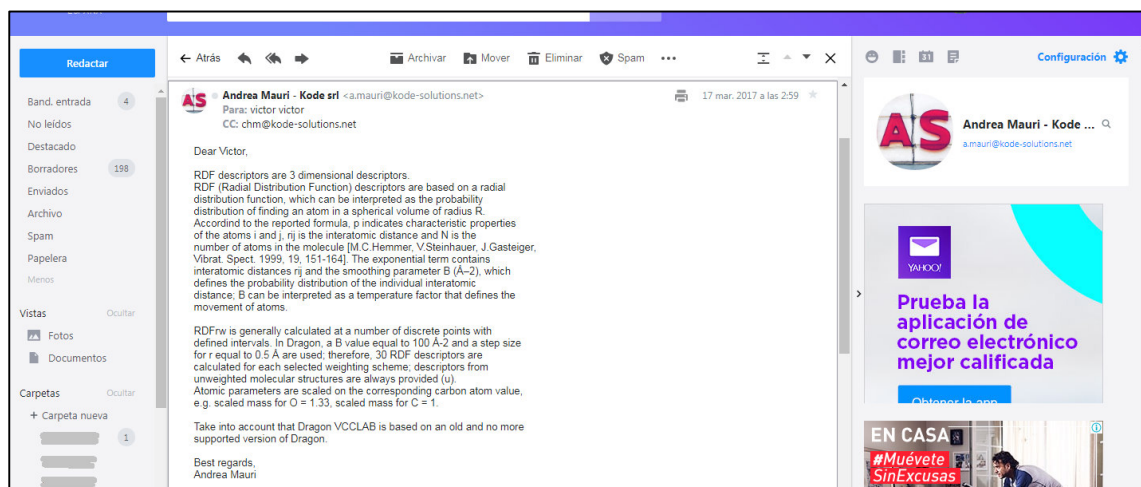
- 28.- Mackintosh CE, Maldonado JA, Ikonomou MG, Gobas FAPC (2006) Sorption of phthalate esters and PCBs in a marine ecosystem. *Environ Sci Technol* 40(11):3481–3488.
- 29.- Matsumoto M, Hirata-Koizumi M, Ema M (2008) Potential adverse effects of phthalic acid esters on human health: a review of recent studies on reproduction. *Regul Toxicol Pharm* 50:37–49
- 30.- Mauri Andrea, Consonni Viviana, Pavan Manuela, and Todeschini Roberto. Dragon software: an easy approach to molecular descriptor calculations. [*match Commun. Math. Comput. Chem.* 56 (2006) 237-248]
- 31.- Mersiowsky Ivo, Weller Martin and Jorgen. Fate of plasticised pvc products under landfill conditions: a laboratory-scale landfill simulation reactor study. *wat. res.* vol. 35, no. 13, pp. 3063–3070, 2001
- 32.- Ojha P.K., Mitra I., Das R.N., Roy K., Further exploring r^2 metrics for validation of QSPR models, *Chemom. Intell. Lab. Syst.* 107, 2011, pp.194–205.
- 33.- Pflughaupt Robin. The epistemological autonomy of chemistry. Master of Arts – Integrated Studies. Athabasca University, Alberta. Canada, 2009.
- 34.- Pramanik Subrata, Roy Kunal. Environmental toxicological fate prediction of diverse organic chemicals based on steady-state compartmental chemical mass ratio using quantitative structure-fate relationship (QSFR). *Chemosphere* 92, 2013, pp. 600-607.
- 35.- Rasmussen, N.C. "Reactor safety study. An assessment of accident risks in U. S. commercial nuclear power plants", WASH-1400 (NUREG-75/014), 1975.
- 36.- Rohac Vladislav, Ruzicka Kvetoslav, Ruzicka Vlastimil, Zaitsau Dzmitry H., Kabo Gennady J., Diky Vladimir, Aim Karel. Vapour pressure of diethyl phthalate. *J. Chem. Thermodynamics* 36, 2004, pp. 929–937.

- 37.- Sacan MT, Ozkul M, Erdem SS (2005) Physico-chemical properties of PCDD/PCDFs and phthalate esters. SAR QSAR Environ Res 16, pp. 443-459.
- 38.- Sadtler Index; *Lide, D.R., G.W.A. Milne (eds.). Handbook of Data on Organic Compounds. Vol. I. 3^a Ed. CRC Press, Inc. Boca Raton, FL. 1994, V1, p. 814.*
- 39.- Schuurmann, G. et. al., Environ Sci Technol, 40, 2006, pp. 7005-11.
- 40.- Stanley, M.K., Robillard, K.A., Staples, C.A., 2003. In: Staples CA (ed) The Handbook of Environmental Chem., Vol. 3, Part Q. Springer, Berlin, pp. 1–7.
- 41.- Staples, Charles A., Peterson D. R., Parkerton, T. F. and Adams, W. J., The Environmental Fate of Phthalate Esters. a Literature Review. Chemosphere, Vol. 35, N°4, 1997, pp. 667-749.
- 42.- Swann RL, Laskowski DA, McCall PJ, *et al.* A rapid method for the estimation of the environmental parameters octanol/water partition coefficient, soil sorption constant, water to air ratio and water solubility, Res. Rev., 85, 1983, pp.17-28.
- 43.- Tienpont Bart. Determination of Phthalates in Environmental, Food and Biomatrices - An Analytical Challenge. February 2004. Doctoral Thesis. Ghent University.
- 44.- Thomsen M, Rasmussen AG, Carlsen L. SAR/QSAR approaches to solubility, partitioning and sorption of phthalates. Chemosphere 38: 2613-2624. (1999). [26]
- 45.- Todeschini R.; Consonni, V.; in Manhold, R.; Kubinyi, H.; Temmerman, H. (Series editors), Handbook of molecular descriptors, Weinheim, New York, Wiley-VCH, (2000).
- 46.- Tropsha A., Gramatica P., Gombar V.K., The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, QSAR Comb. Sci. 22 (2003) 69–77.

- 47.- US Environmental Protection Agency National Primary Drinking Water Regulations; Fed. Reg., Part 12, 40 CFR Part 141, p.395. Washington, DC, 1 July 1991
- 48.- U.S. Environmental Protection Agency – USA Phathales Action Plan. Revised 03/14/2012.
- 49.- US Environmental Protection Agency; Estimation Program Interface (EPI) Suite. Ver. 141. Nov, 2012. Available from, as of Feb 12, 2015: <http://www.epa.gov/oppt/exposure/pubs/episuitedi.htm>
- 50.- Villaveces Cardoso José Luis. Química y Epistemología, una relación esquivada., Revista Colombiana de Filosofía de la Ciencia, año 1, Nos 2 y 3, pag 9-26, 2000.
- 51.- Vracko M. and Gasteiger J., A QSAR Study on a Set of 105 Flavonoid Derivatives Using Descriptors Derived From 3D Structures, Internet Electron. J. Mol. Des. 2002, 1, 527–544,
- 52.- Weininger David , SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.*, 1988, 28 (1), pp 31–36)
- 53.- Wypych George. Handbook of Platicizers.. ChemTec Publishing. Toronto – New York 2004
- 54.- Yousefinejad S., Hemmateenejad B., Chemometrics tools in QSAR/QSPR studies: A historical perspective, Chemometr. Intell. Lab. Syst. 2015
- 55.- Zhang S., Golbraikh A., Oloff S., Kohh H., Tropsha A., A novel automated lazy learning QSAR (ALL-QSAR) approach:method development, applications, and virtual screening of chemical databases using validated ALL-QSAR models, *J. Chem. Inf. Model.* 46(2006) 1984–1995.

ANEXOS

ANEXO 1

1.a Comunicación VCCLAB1b Comunicación Kode-Solutions

Año	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH	BI	BJ	BK	BL	BM	BN	BO	BP	BQ	BR	BS	BT	BU	BV	BW	BX	BY	BZ	CA	CB	CC	CD	CE	CF	CG	CH	CI	CJ	CK	CL	CM	CN	CO	CP	CQ	CR	CS	CT	CU	CV	CW	CX	CY	CZ	DA	DB	DC	DD	DE	DF	DG	DH	DI	DJ	DK	DL	DM	DN	DO	DP	DQ	DR	DS	DT	DU	DV	DW	DX	DY	DZ	EA	EB	EC	ED	EE	EF	EG	EH	EI	EJ	EK	EL	EM	EN	EO	EP	EQ	ER	ES	ET	EU	EV	EW	EX	EY	EZ	FA	FB	FC	FD	FE	FF	FG	FH	FI	FJ	FK	FL	FM	FN	FO	FP	FQ	FR	FS	FT	FU	FV	FW	FX	FY	FZ	GA	GB	GC	GD	GE	GF	GG	GH	GI	GJ	GK	GL	GM	GN	GO	GP	GQ	GR	GS	GT	GU	GV	GW	GX	GY	GZ	HA	HB	HC	HD	HE	HF	HG	HH	HI	HJ	HK	HL	HM	HN	HO	HP	HQ	HR	HS	HT	HU	HV	HW	HX	HY	HZ	IA	IB	IC	ID	IE	IF	IG	IH	II	IJ	IK	IL	IM	IN	IO	IP	IQ	IR	IS	IT	IU	IV	IW	IX	IY	IZ	JA	JB	JC	JD	JE	JF	JG	JH	JI	JJ	JK	JL	JM	JN	JO	JP	JQ	JR	JS	JT	JU	JV	JW	JX	JY	JZ	KA	KB	KC	KD	KE	KF	KG	KH	KI	KJ	KK	KL	KM	KN	KO	KP	KQ	KR	KS	KT	KU	KV	KW	KX	KY	KZ	LA	LB	LC	LD	LE	LF	LG	LH	LI	LJ	LK	LL	LM	LN	LO	LP	LQ	LR	LS	LT	LU	LV	LW	LX	LY	LZ	MA	MB	MC	MD	ME	MF	MG	MH	MI	MJ	MK	ML	MM	MN	MO	MP	MQ	MR	MS	MT	MU	MV	MW	MX	MY	MZ	NA	NB	NC	ND	NE	NF	NG	NH	NI	NJ	NK	NL	NM	NO	NP	NQ	NR	NS	NT	NU	NV	NW	NX	NY	NZ	OA	OB	OC	OD	OE	OF	OG	OH	OI	OJ	OK	OL	OM	ON	OO	OP	OQ	OR	OS	OT	OU	OV	OW	OX	OY	OZ	PA	PB	PC	PD	PE	PF	PG	PH	PI	PJ	PK	PL	PM	PN	PO	PP	PQ	PR	PS	PT	PU	PV	PW	PX	PY	PZ	QA	QB	QC	QD	QE	QF	QG	QH	QI	QJ	QK	QL	QM	QN	QO	QP	QQ	QR	QS	QT	QU	QV	QW	QX	QY	QZ	RA	RB	RC	RD	RE	RF	RG	RH	RI	RJ	RK	RL	RM	RN	RO	RP	RQ	RR	RS	RT	RU	RV	RW	RX	RY	RZ	SA	SB	SC	SD	SE	SF	SG	SH	SI	SJ	SK	SL	SM	SN	SO	SP	SQ	SR	SS	ST	SU	SV	SW	SX	SY	SZ	TA	TB	TC	TD	TE	TF	TG	TH	TI	TJ	TK	TL	TM	TN	TO	TP	TQ	TR	TS	TT	TU	TV	TW	TX	TY	TZ	UA	UB	UC	UD	UE	UF	UG	UH	UI	UJ	UK	UL	UM	UN	UO	UP	UQ	UR	US	UT	UU	UV	UW	UX	UY	UZ	VA	VB	VC	VD	VE	VF	VG	VH	VI	VJ	VK	VL	VM	VN	VO	VP	VQ	VR	VS	VT	VU	VV	VW	VX	VY	VZ	WA	WB	WC	WD	WE	WF	WG	WH	WI	WJ	WK	WL	WM	WN	WO	WP	WQ	WR	WS	WT	WU	WV	WW	WX	WY	WZ	XA	XB	XC	XD	XE	XF	XG	XH	XI	XJ	XK	XL	XM	XN	XO	XP	XQ	XR	XS	XT	XU	XV	XW	XY	XZ	YA	YB	YC	YD	YE	YF	YG	YH	YI	YJ	YK	YL	YM	YN	YO	YP	YQ	YR	YS	YT	YU	YV	YW	YZ	ZA	ZB	ZC	ZD	ZE	ZF	ZG	ZH	ZI	ZJ	ZK	ZL	ZM	ZN	ZO	ZP	ZQ	ZR	ZS	ZT	ZU	ZV	ZW	ZX	ZY	ZZ																																																																																																																	
Paño	14530	14531	14532	14533	14534	14535	14536	14537	14538	14539	14540	14541	14542	14543	14544	14545	14546	14547	14548	14549	14550	14551	14552	14553	14554	14555	14556	14557	14558	14559	14560	14561	14562	14563	14564	14565	14566	14567	14568	14569	14570	14571	14572	14573	14574	14575	14576	14577	14578	14579	14580	14581	14582	14583	14584	14585	14586	14587	14588	14589	14590	14591	14592	14593	14594	14595	14596	14597	14598	14599	14600	14601	14602	14603	14604	14605	14606	14607	14608	14609	14610	14611	14612	14613	14614	14615	14616	14617	14618	14619	14620	14621	14622	14623	14624	14625	14626	14627	14628	14629	14630	14631	14632	14633	14634	14635	14636	14637	14638	14639	14640	14641	14642	14643	14644	14645	14646	14647	14648	14649	14650	14651	14652	14653	14654	14655	14656	14657	14658	14659	14660	14661	14662	14663	14664	14665	14666	14667	14668	14669	14670	14671	14672	14673	14674	14675	14676	14677	14678	14679	14680	14681	14682	14683	14684	14685	14686	14687	14688	14689	14690	14691	14692	14693	14694	14695	14696	14697	14698	14699	14700	14701	14702	14703	14704	14705	14706	14707	14708	14709	14710	14711	14712	14713	14714	14715	14716	14717	14718	14719	14720	14721	14722	14723	14724	14725	14726	14727	14728	14729	14730	14731	14732	14733	14734	14735	14736	14737	14738	14739	14740	14741	14742	14743	14744	14745	14746	14747	14748	14749	14750	14751	14752	14753	14754	14755	14756	14757	14758	14759	14760	14761	14762	14763	14764	14765	14766	14767	14768	14769	14770	14771	14772	14773	14774	14775	14776	14777	14778	14779	14780	14781	14782	14783	14784	14785	14786	14787	14788	14789	14790	14791	14792	14793	14794	14795	14796	14797	14798	14799	14800	14801	14802	14803	14804	14805	14806	14807	14808	14809	14810	14811	14812	14813	14814	14815	14816	14817	14818	14819	14820	14821	14822	14823	14824	14825	14826	14827	14828	14829	14830	14831	14832	14833	14834	14835	14836	14837	14838	14839	14840	14841	14842	14843	14844	14845	14846	14847	14848	14849	14850	14851	14852	14853	14854	14855	14856	14857	14858	14859	14860	14861	14862	14863	14864	14865	14866	14867	14868	14869	14870	14871	14872	14873	14874	14875	14876	14877	14878	14879	14880	14881	14882	14883	14884	14885	14886	14887	14888	14889	14890	14891	14892	14893	14894	14895	14896	14897	14898	14899	14900	14901	14902	14903	14904	14905	14906	14907	14908	14909	14910	14911	14912	14913	14914	14915	14916	14917	14918	14919	14920	14921	14922	14923	14924	14925	14926	14927	14928	14929	14930	14931	14932	14933	14934	14935	14936	14937	14938	14939	14940	14941	14942	14943	14944	14945	14946	14947	14948	14949	14950	14951	14952	14953	14954	14955	14956	14957	14958	14959	14960	14961	14962	14963	14964	14965	14966	14967	14968	14969	14970	14971	14972	14973	14974	14975	14976	14977	14978	14979	14980	14981	14982	14983	14984	14985	14986	14987	14988	14989	14990	14991	14992	14993	14994	14995	14996	14997	14998	14999	15000	15001	15002	15003	15004	15005	15006	15007	15008	15009	15010	15011	15012	15013	15014	15015	15016	15017	15018	15019	15020	15021	15022	15023	15024	15025	15026	15027	15028	15029	15030	15031	15032	15033	15034	15035	15036	15037	15038	15039	15040	15041	15042	15043	15044	15045	15046	15047	15048	15049	15050	15051	15052	15053	15054	15055	15056	15057	15058	15059	15060	15061	15062	15063	15064	15065	15066	15067	15068	15069	15070	15071	15072	15073	15074	15075	15076	15077	15078	15079	15080	15081	15082	15083	15084	15085	15086	15087	15088	15089	15090	15091	15092	15093	15094	15095	15096	15097	15098	15099	15100	15101	15102	15103	15104	15105	15106	15107	15108	15109	15110	15111	15112	15113	15114	15115	15116	15117	15118	15119	15120	15121	15122	15123	15124	15125	15126	15127	15128	15129	15130	15131	15132	15133	15134	15135	15136	15137	15138	15139	15140	15141	15142	15143	15144	15145	15146	15147	15148	15149	15150	15151	15152	15153	15154	15155	15156	15157	15158	15159	15160	15161	15162	15163	15164	15165	15166	15167	15168	15169	15170	15171	15172	15173	15174	15175	15176	15177	15178	15179	15180	15181	15182	15183	15184	15185	15186	15187	15188	15189	15190	15191	15192	15193	15194	15195	15196	15197	15198	15199	15200	15201	15202	15203	15204	15205	15206	15207	15208	15209	15210	15211	15212	15213	15214	15215	15216	15217	15218	15219	15220	15221	15222	15223	15224	15225	15226	15227	15228	15229	15230	15231	15232	15233	15234	15235	15236	15237	15238	15239	15240	15241	15242	15243	15244	15245	15246	15247	15248	15249	15250	15251	15252	15253	15254	15255	15256	15257	15258	15259	15260	15261	15262	15263	15264	15265	15266	15267	15268	15269	15270	15271	15272	15273	15274	15275	15276	15277	15278	15279	15280	15281	15282	15283	15284	15285	15286	15287	15288	15289	15290	15291	15292	15293	15294	15295	15296	15297	15298	15299	15300	15301	15302	15303	15304	15305	15306	15307	15308	15309	15310	15311	15312	15313	15314	15315	15316	15317	15318	15319	15320	15321	15322	15323	15324	15325	15326	15327	15328	15329	15330	15331	15332	15333	15334	15335	15336	15337	15338	15339</

	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H				
Atomo	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66					
Radio	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500				
tomo 48 con	2.46	6.8	6.81	6.53	5.64	7.28	7.040	7.64	8.42	8.85	8.76	8.49	8.68	10.81	10.66	12.12	10.78	10.93					
Je distancias	12.040	7.700	7.690	7.970	8.860	7.220	7.460	6.860	6.080	5.650	5.740	6.010	5.820	3.690	3.840	2.380	3.720	3.570					
la diferencia	144.962	59.290	59.136	63.521	78.500	52.128	55.652	47.060	36.966	31.923	32.948	36.120	33.872	13.616	14.746	5.664	13.838	12.745					
0 (producto)	-14496.160	-5929.000	-5913.610	-6352.090	-7849.960	-5212.840	-5565.160	-4705.960	-3696.640	-3192.250	-3294.760	-3612.010	-3387.240	-1361.610	-1474.560	-566.440	-1383.840	-1274.490					
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00		
p 1	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083					
p 2	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083					
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00		
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264					
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00		
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944					
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00		
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388					
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00		

	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H				
Atomo	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66					
Radio	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500					
tomo 49 con	4.48	4.49	4.14	3.39	4.99	4.940	5.61	6.15	6.43	6.31	6.14	6.35	8.39	8.23	9.78	8.51	8.66					
Je distancias	10.020	10.010	10.360	11.110	9.510	9.560	8.890	8.350	8.070	8.190	8.360	8.150	6.110	6.270	4.720	5.990	5.840					
la diferencia	100.400	100.200	107.330	123.432	90.440	91.394	79.032	69.723	65.125	67.076	69.890	66.423	37.332	39.313	22.278	35.880	34.106					
0 (producto)	-10040.040	-10020.010	-10732.960	-12343.210	-9044.010	-9139.360	-7903.210	-6972.250	-6512.490	-6707.610	-6988.960	-6642.250	-3733.210	-3931.290	-2227.840	-3588.010	-3410.560					
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00		
p 1	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083					
p 2	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083					
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00		
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264					
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00		
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944					
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00		
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388					
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00		

	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H	H				
Atomo	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66					
Radio	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500					
tomo 50 con	1.78	3.05	2.92	2.52	4.650	4.78	4.39	3.72	2.49	3.07	4.17	5.22	4.39	6.49	5.39	6.08					
Je distancias	12.720	11.450	11.580	11.980	9.850	9.720	10.110	10.780	12.010	11.430	10.330	9.280	10.110	8.010	9.110	8.420					
la diferencia	161.798	131.103	134.096	143.520	97.023	94.478	102.212	116.208	144.240	130.645	106.709	86.118	102.212	64.160	82.992	70.896					
0 (producto)	-16179.840	-13110.250	-13409.640	-14352.040	-9702.250	-9447.840	-10221.210	-11620.840	-14424.010	-13064.490	-10670.890	-8611.840	-10221.210	-6416.010	-8299.210	-7089.640					
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00		
p 1	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083					
p 2	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083					
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00		
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264					
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00		
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944					
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00		
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388					
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00		

	H	H	H	H	H	H	H	H	H	H	H	H			
Atomo	55	56	57	58	59	60	61	62	63	64	65	66			
Radio	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500			
tomo 54 con	3.05	2.460	2.5	2.89	2.52	4.4	4.64	5.19	4.98	7.22	6.65	6.81			
Je distancias	11.450	12.040	12.000	11.610	11.980	10.100	9.860	9.310	9.520	7.280	7.850	7.690			
la diferencia	131.103	144.962	144.000	134.792	143.520	102.010	97.220	86.676	90.630	52.998	61.623	59.136			
0 (producto)	-13110.250	-14496.160	-14400.000	-13479.210	-14352.040	-10201.000	-9721.960	-8667.610	-9063.040	-5299.840	-6162.250	-5913.610			
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00
p 1	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
p 2	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00

	H	H	H	H	H	H	H	H	H	H	H	H			
Atomo	56	57	58	59	60	61	62	63	64	65	66				
Radio	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500			
tomo 55 con	1.78	1.780	3.03	4.18	5.11	4.26	5.36	6.06	7.58	7.19	6.61				
Je distancias	12.720	12.720	11.470	10.320	9.390	10.240	9.140	8.440	6.920	7.310	7.890				
la diferencia	161.798	161.798	131.561	106.502	88.172	104.858	83.540	71.234	47.886	53.436	62.252				
0 (producto)	-16179.840	-16179.840	-13156.090	-10650.240	-8817.210	-10485.760	-8353.960	-7123.360	-4788.640	-5343.610	-6225.210				
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00
p 1	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
p 2	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00

	H	H	H	H	H	H	H	H	H	H	H				
Atomo	57	58	59	60	61	62	63	64	65	66					
Radio	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500				
tomo 56 con	1.78	3.490	4.27	5.94	5.42	5.96	6.43	8.32	8.04	7.67					
Je distancias	12.720	11.010	10.230	8.560	9.080	8.540	8.070	6.180	6.460	6.830					
la diferencia	161.798	121.220	104.653	73.274	82.446	72.932	65.125	38.192	41.732	46.649					
0 (producto)	-16179.840	-12122.010	-10465.290	-7327.360	-8244.640	-7293.160	-6512.490	-3819.240	-4173.160	-4664.890					
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00	
p 1	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
p 2	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00

	H	H	H	H	H	H	H	H	H			
Atomo	58	59	60	61	62	63	64	65	66			
Radio	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500			
tomo 57 con	1.85	3.090	4.67	3.96	4.22	4.89	6.62	6.51	6.02			
de distancias	12.650	11.410	9.830	10.540	10.280	9.610	7.880	7.990	8.480			
la diferencia	160.023	130.188	96.629	111.092	105.678	92.352	62.094	63.840	71.910			
0 (producto)	-16002.250	-13018.810	-9662.890	-11109.160	-10567.840	-9235.210	-6209.440	-6384.010	-7191.040			
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00
p 1	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
p 2	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00

	H	H	H	H	H	H	H	H	H			
Atomo	59	60	61	62	63	64	65	66				
Radio	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500			
tomo 58 con	1.78	3.050	2.46	2.53	3.11	4.83	4.68	4.32				
de distancias	12.720	11.450	12.040	11.970	11.390	9.670	9.820	10.180				
la diferencia	161.798	131.103	144.962	143.281	129.732	93.509	96.432	103.632				
0 (producto)	-16179.840	-13110.250	-14496.160	-14328.090	-12973.210	-9350.890	-9643.240	-10363.240				
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00
p 1	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
p 2	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00

	H	H	H	H	H	H	H	H				
Atomo	60	61	62	63	64	65	66					
Radio	14.500	14.500	14.500	14.500	14.500	14.500	14.500	14.500				
tomo 59 con	2.5	3.050	3.03	2.47	4.78	4.31	4.65					
de distancias	12.000	11.450	11.470	12.030	9.720	10.190	9.850					
la diferencia	144.000	131.103	131.561	144.721	94.478	103.836	97.023					
0 (producto)	-14400.000	-13110.250	-13156.090	-14472.090	-9447.840	-10383.610	-9702.250					
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00
p 1	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
p 2	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083			
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264	0.264			
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944	0.944			
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388	0.388			
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00

	H	H	H	H	H	H			
Atomo	61	62	63	64	65	66			
Radio	14.500	14.500	14.500	14.500	14.500	14.500			
tomo 60 con	1.78	3.050	2.48	3.72	2.5	3.07			
le distancias	12.720	11.450	12.020	10.780	12.000	11.430			
la diferencia	161.798	131.103	144.480	116.208	144.000	130.645			
O (producto)	-16179.840	-13110.250	-14448.040	-11620.840	-14400.000	-13064.490			
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00
ρ 1	0.083	0.083	0.083	0.083	0.083	0.083			
ρ 2	0.083	0.083	0.083	0.083	0.083	0.083			
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00
	0.264	0.264	0.264	0.264	0.264	0.264			
	0.264	0.264	0.264	0.264	0.264	0.264			
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00
	0.944	0.944	0.944	0.944	0.944	0.944			
	0.944	0.944	0.944	0.944	0.944	0.944			
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00
	0.388	0.388	0.388	0.388	0.388	0.388			
	0.388	0.388	0.388	0.388	0.388	0.388			
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00

	H	H	H	H	H				
Atomo	62	63	64	65	66				
Radio	14.500	14.500	14.500	14.500	14.500				
tomo 61 con	2.48	3.050	3.72	3.07	2.5				
le distancias	12.020	11.450	10.780	11.430	12.000				
la diferencia	144.480	131.103	116.208	130.645	144.000				
O (producto)	-14448.040	-13110.250	-11620.840	-13064.490	-14400.000				
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00
ρ 1	0.083	0.083	0.083	0.083	0.083				
ρ 2	0.083	0.083	0.083	0.083	0.083				
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00
	0.264	0.264	0.264	0.264	0.264				
	0.264	0.264	0.264	0.264	0.264				
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00
	0.944	0.944	0.944	0.944	0.944				
	0.944	0.944	0.944	0.944	0.944				
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00
	0.388	0.388	0.388	0.388	0.388				
	0.388	0.388	0.388	0.388	0.388				
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00

	H	H	H	H				
Atomo	63	64	65	66				
Radio	14.500	14.500	14.500	14.500				
tomo 62 con	1.78	2.480	3.05	2.48				
le distancias	12.720	12.020	11.450	12.020				
la diferencia	161.798	144.480	131.103	144.480				
O (producto)	-16179.840	-14448.040	-13110.250	-14448.040				
Exponencial	0.000	0.000	0.000	0.000	0.000	0.000	RDFu	0.00
ρ 1	0.083	0.083	0.083	0.083				
ρ 2	0.083	0.083	0.083	0.083				
RDFm	0.000	0.000	0.000	0.000	0.000	0.000	RDFm	0.00
	0.264	0.264	0.264	0.264				
	0.264	0.264	0.264	0.264				
RDFv	0.000	0.000	0.000	0.000	0.000	0.000	RDFv	0.00
	0.944	0.944	0.944	0.944				
	0.944	0.944	0.944	0.944				
RDFe	0.000	0.000	0.000	0.000	0.000	0.000	RDFe	0.00
	0.388	0.388	0.388	0.388				
	0.388	0.388	0.388	0.388				
RDFp	0.000	0.000	0.000	0.000	0.000	0.000	RDFp	0.00

	H	H	H			
Atomo	64	65	66			
Radio	14.500	14.500	14.500			
tomo 63 con	2.48	2.480	3.05			
le distancias	12.020	12.020	11.450			
la diferencia	144.480	144.480	131.103			
O (producto)	-14448.040	-14448.040	-13110.250			
Exponencial	0.000	0.000	0.000	0.000	RDFu	0.00
p_1	0.083	0.083	0.083			
p_2	0.083	0.083	0.083			
RDFm	0.000	0.000	0.000	0.000	RDFm	0.00
	0.264	0.264	0.264			
	0.264	0.264	0.264			
RDFv	0.000	0.000	0.000	0.000	RDFv	0.00
	0.944	0.944	0.944			
	0.944	0.944	0.944			
RDFe	0.000	0.000	0.000	0.000	RDFe	0.00
	0.388	0.388	0.388			
	0.388	0.388	0.388			
RDFp	0.000	0.000	0.000	0.000	RDFp	0.00

	H	H				
Atomo	65	66				
Radio	14.500	14.500				
tomo 64 con	1.78	1.780				
le distancias	12.720	12.720				
la diferencia	161.798	161.798				
O (producto)	-16179.840	-16179.840				
Exponencial	0.000	0.000	0.000	RDFu	0.00	
p_1	0.083	0.083				
p_2	0.083	0.083				
RDFm	0.000	0.000	0.000	RDFm	0.00	
	0.264	0.264				
	0.264	0.264				
RDFv	0.000	0.000	0.000	RDFv	0.00	
	0.944	0.944				
	0.944	0.944				
RDFe	0.000	0.000	0.000	RDFe	0.00	
	0.388	0.388				
	0.388	0.388				
RDFp	0.000	0.000	0.000	RDFp	0.00	

	H				
Atomo	66				
Radio	14.500				
tomo 65 con	1.78				
le distancias	12.720				
la diferencia	161.798				
O (producto)	-16179.840				
Exponencial	0.000	0.000	RDFu	0.00	
p_1	0.083				
p_2	0.083				
RDFm	0.000	0.000	RDFm	0.00	
	0.264				
	0.264				
RDFv	0.000	0.000	RDFv	0.00	
	0.944				
	0.944				
RDFe	0.000	0.000	RDFe	0.00	
	0.388				
	0.388				
RDFp	0.000	0.000	RDFp	0.00	

ANEXO 3

DESCRIPTOR RDFv

r[Å]	DMP	DAP	DBP	BBP	DEP
1	1.439	2.118	3.034	2.891	1.971
1.5	4.633	6.677	10.116	9.86	6.46
2	0.793	1.93	1.376	1.554	0.927
2.5	9.263	11.859	16.076	17.408	10.646
3	1.698	2.373	2.969	2.348	1.97
3.5	1.903	3.44	2.393	2.87	2.009
4	2.005	3.086	4.153	3.975	3.169
4.5	2.304	1.951	2.509	5.572	2.986
5	4.67	7.445	11.557	9.768	7.3
5.5	0.427	0.774	0.534	3.336	0.606
6	2.87	6.288	4.828	4.928	4.651
6.5	3.124	6.04	7.773	7.251	5.295
7	1.54	2.433	2.764	4.749	1.671
7.5	0.808	4.714	8.733	4.576	2.872
8	0.123	1.904	2.494	7.474	1.268
8.5	0.077	4.111	6.649	7.327	1.949
9	0	1.337	3.762	6.58	0.301
9.5	0	1.163	2.329	4.087	0.399
10	0	1.267	4.051	3.694	0.496
10.5	0	0.953	2.368	2.38	0.203
11	0	0.857	0.805	1.685	0
11.5	0	0.118	2.273	1.654	0
12	0	0.296	2.175	2.778	0
12.5	0	0.049	1.068	1.226	0
13	0	0.088	0.255	1.37	0
13.5	0	0	0.943	0.333	0
14	0	0	0.765	0.984	0
14.5	0	0	0.179	0.531	0
15	0	0	0	0.204	0
15.5	0	0	0	0	0

Fuente. E-Dragon 1.0

ANEXO 4
DESCRIPTOR RDFe

r[Å]	DMP	DAP	DBP	BBP	DEP
1	4.527	6.669	9.555	9.105	6.203
1.5	6.087	8.13	11.573	11.315	7.916
2	4.741	7.802	5.876	6.413	4.441
2.5	16.242	21.499	43.007	38.737	25.004
3	4.14	7.808	16.739	10.451	6.838
3.5	6.742	11.034	8.14	9.682	7.245
4	6.132	10.629	14.408	12.948	9.718
4.5	8.138	7.306	12.012	15.975	13.297
5	14.266	20.649	36.482	29.858	19.792
5.5	2.315	3.969	2.606	7.984	3.317
6	9.39	16.586	13.842	14.87	13.538
6.5	5.545	12.187	15.867	14.752	9.21
7	7.895	6.47	9.612	14.613	5.701
7.5	2.697	8.683	20.605	12.854	7.113
8	1.224	7.097	8.369	15.569	6.55
8.5	0.762	9.313	15.94	16.587	6.929
9	0	2.958	10.674	11.997	2.446
9.5	0	4.871	8.2	10.601	2.409
10	0	3.618	12.684	11.757	1.567
10.5	0	1.653	10.714	11.247	2.015
333311	0	3.014	4.4	4.251	0
11.5	0	0.859	4.291	3.489	0
12	0	1.514	7.079	6.232	0
12.5	0	0.483	5.46	3.406	0
13	0	0.876	2.529	3.908	0
13.5	0	0	1.244	1.928	0
14	0	0	2.431	2.869	0
14.5	0	0	1.773	2.964	0
15	0	0	0	1.941	0
15.5	0	0	0	0.001	0

Fuente. E-Dragon 1.0

ANEXO 5
DESCRIPTOR RDFu

r[Å]	DNHP	DNOP	610P	DIOP	DEHP	BOHP
1	13.696	17.254	17.254	17.254	17.254	13.696
1.5	14.648	18.305	18.305	18.306	18.306	14.648
2	7.133	8.655	8.655	9.352	9.743	8.124
2.5	63.724	83.51	83.508	79.93	78.854	59.823
3	28.578	39.737	39.729	41.337	41.767	29.616
3.5	7.081	7.081	7.081	15.437	20.328	11.677
4	16.171	18.98	18.978	23.324	28.92	24.486
4.5	11.813	12.551	12.533	17.433	27.873	17.572
5	49.502	65.805	65.843	57.952	53.181	46.044
5.5	2.567	2.743	2.741	8.365	19.2	9.941
6	15.257	15.25	15.237	22.264	31.589	21.625
6.5	26.689	45.324	45.326	35.265	48.156	30.915
7	8.549	9.142	9.147	10.774	28.621	18.479
7.5	35.933	57.281	57.29	44.252	31.108	31.703
8	8.471	8.473	8.477	9.757	23.644	14.245
8.5	18.522	22.986	20.769	23.886	40.699	26.754
9	18.291	27.046	31.008	20.618	29.959	19.477
9.5	13.861	11.981	12.922	14.16	30.296	13.628
10	23.164	37.098	42.687	34.338	31.144	19.823
10.5	13.273	13.448	12.538	16.558	25.902	14.809
11	15.424	18.109	20.765	25.957	34.208	14.281
11.5	14.458	24.008	20.881	22.515	23.543	9.606
12	18.599	24.694	21.579	29.765	18.629	12.092
12.5	15.381	23.134	23.328	23.651	22.213	12.056
13	12.13	15.801	14.944	21.424	16.649	7.896
13.5	5.548	16.757	13.97	21.027	8.973	3.057
14	10.289	18.028	16.223	17.403	12.337	6.528
14.5	7.518	20.798	18.33	19.36	10.532	5.332
15	5.118	13.191	13.541	13.816	5.293	2.509
15.5	3.529	12.462	11.912	12.583	4.062	0.667

Fuente. E-Dragon 1.0

ANEXO 6
DESCRIPTOR RDFm

r[Å]	DNOP	610P	DIOP	BOP	DEHP	DHP
1	1.476	1.476	1.476	1.177	1.476	1.177
1.5	18.889	18.889	18.889	15.233	18.889	15.233
2	0.797	0.797	0.773	0.767	0.88	0.672
2.5	25.586	25.585	27.667	22.933	28.111	21.467
3	3.993	3.992	5.94	6.894	9.993	3.913
3.5	3.947	3.947	4.538	4.056	5.008	3.947
4	5.412	5.409	5.419	4.951	5.446	4.784
4.5	1.998	1.997	2.678	2.858	3.737	1.993
5	22.296	22.299	20.321	17.006	19.364	17.88
5.5	0.301	0.302	0.696	1.054	2.226	0.3
6	4.794	4.779	4.946	8.147	8.554	4.794
6.5	10.179	10.172	9.383	10.214	15.022	8.734
99997	1.263	1.273	1.468	2.632	9.111	1.259
7.5	17.883	17.895	15.496	10.651	11.303	13.058
8	1.79	1.792	1.819	2.002	3.858	1.79
8.5	6.822	6.783	6.933	7.645	12.001	6.719
9	4.311	4.48	4.869	4.304	6.681	3.854
9.5	2.881	2.968	3.058	1.534	4.242	3.055
10	12.211	11.932	13.676	4.821	6.531	6.576
10.5	1.531	1.454	2.215	1.135	4.126	1.461
11	5.429	5.334	6.92	3.838	6.444	3.885
11.5	5.059	4.592	5.388	2.733	5.605	3.421
12	5.7	4.229	7.856	2.089	3.004	3.269
12.5	4.205	4.779	4.088	1.47	2.107	1.259
13	1.255	1.305	1.861	0.674	2.069	0.995
13.5	3.846	3.876	4.085	0.915	1.271	1.163
14	2.238	3.063	1.681	0.423	1.313	1.269
14.5	3.934	4.901	1.589	0.594	0.635	0.724
15	2.379	2.005	2.358	0.081	0.569	0.179
15.5	1.294	1.502	1.201	0.064	0.96	0.858

Fuente. E-Dragon 1.0