

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

UNIVERSITE BADJI MOKHTAR - ANNABA
BADJI MOKHTAR – ANNABA UNIVERSITY



جامعة باجي مختار – عنابة

Faculté : Science
Département : Chimie
Domaine :
Filière : Chimie
Spécialité : Electrochimie

Mémoire

Présenté en vue de l'obtention du Diplôme de Master
Thème:

**Prédiction d'une propriété électrochimique des composés
Organique par l'approche QSPR**

Présenté par : *Chaab Mourad*

Rouabhia Mohamed

Encadrant : *Ziani Nadia*

Grade *CMB*

Université Badji Mokhtar annaba

Jury de Soutenance :

Dadda Naouri	Dr	<i>Université Badji Mokhtar annaba</i>	Président
Ziani Nadia	MCB	<i>Université Badji Mokhtar annaba</i>	Encadrant
Berrezag Kamel	Dr	<i>Université Badji Mokhtar annaba</i>	Examineur

Année Universitaire : 2022/2023

Remerciement

Remerciement

Notre premier remerciement à " ALLAH " le
tout

puissant de nous avoir offert la force
pour accomplir ce
travail "Dieu merci".

Nous tenons à remercier vivement notre
encadreur Mme Ziani Nadia maitre de
conference

au département de chimie et

Mr Dr Bouaoune
Ahmed, pour sa patience, sa
disponibilité,

sa gentillesse et surtout ses bons
conseils

afin d'élaborer ce mémoire dans de bonnes
conditions.

Nous lui exprimons toute notre
reconnaissance ;

A toute l'équipe pédagogique de
département de chimie;
Aux membres du jury qui s'apprêtent
à évaluer notre travail .



DEDICACES

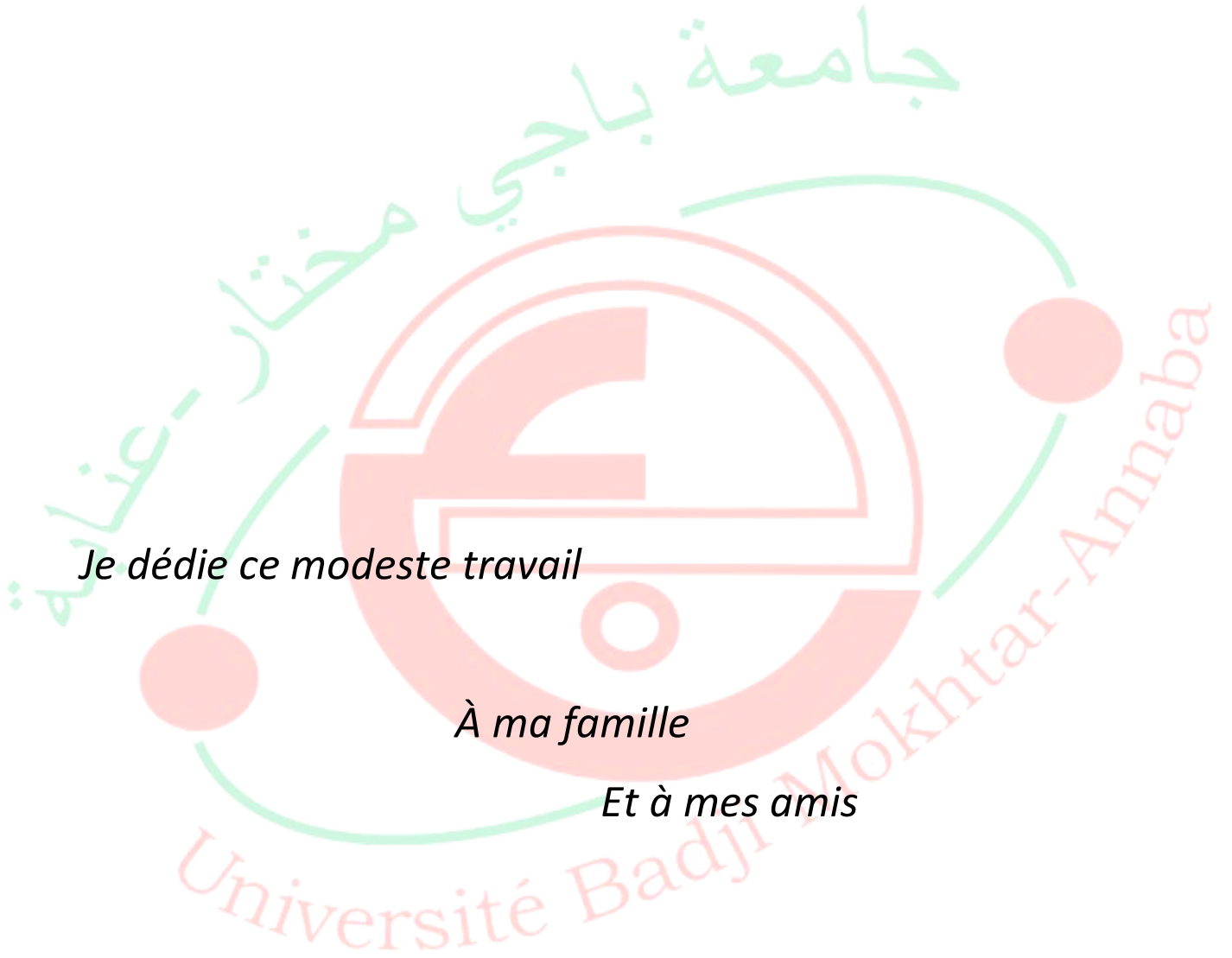
Je dédie ce modeste travail :

*À mes chers famille pour leur soutien , leur patience ,
leur encouragement durant mon parcours scolaire.*

*À tous mes amis , Et l'ensemble des étudiants de
promo LMD De l'année 2022-2023*

Merci d'être toujours là pour moi.

Mohamed



Je dédie ce modeste travail

À ma famille

Et à mes amis

Mourad



ملخص

الهدف من هذا العمل هو تنفيذ اجراء العلاقة الكمية بنية - مميزات(QSPR) تم إجراء تحسين هندسة الجزيئات وحساب الواصفات الجزيئية والاختيار فيما بينها لمجموعة فرعية من المتغيرات التوضيحية ذات الصلة بواسطة الخوارزمية الجينية بواسطة البرنامج المتخصص (MOBDGS). النموذج الذي تم اختياره بواسطة الانحدار الخطي المتعدد مستقر وقوي مع قدرة جيدة إلى حد ما على التنبؤ بخصائص المركبات الجديدة. أن يعالج مشكلة نقص البيانات التجريبية بسهولة QSPR في هذا العمل ، يمكن لنموذج الكلمات المفتاحية مشتقات البروبان - الواصفات الجزيئية - QSPR نموذج - التوتر السطحي

Summary

The objective of this work is to implement the Quantitative Structure-Property Relationship (QSPR) procedure for the prediction of the surface tension of propane derivatives, given their importance in the field of electrochemistry.

The optimization of the geometries of the molecules, the calculation of the molecular descriptors and the choice among them of a subset of relevant explanatory variables by genetic descriptors and the choice among them of a subset of relevant explanatory variables by genetic algorithm were carried out by the specialized software (MOBDGS).

The model selected by multiple linear regression is stable, robust with fairly good ability to predict the properties of new compounds.

In this work, the QSPR model can easily remedy the problem of lack of experimental data.

Keywords: Surface tension – QSPR model – Molecular descriptors – Propane derivatives.

Résumé

L'objectif de ce travail est de mettre en œuvre la procédure des Relations Structure-Propriété Quantitatives (QSPR) pour la prédiction de la tension superficielle des dérivées de propane, vues leurs importances dans le domaine de l'électrochimie.

L'optimisation des géométries des molécules, le calcul des descripteurs moléculaires et le choix d'un sous-ensemble de variables explicatives pertinentes par algorithme génétique ont été effectuées par le logiciel spécialisé (MOBDGS) . Le modèle sélectionné par régression linéaire multiple est stable, robuste avec bonnes capacités à prédire les propriétés de nouveaux composés.

Dans ce travail le modèle QSPR peuvent aisément remédier au problème de carence des données expérimentales.

Mots clés: Dérives du propane - Tension superficielle – Descripteurs moléculaires – Modèle QSPR.

SOMMAIRE

Résumé	1
Liste des tableaux	
Listes des figures	
Symbole et abréviations	
Introduction Générale.....	4
PARTIE GENERALITEE	
I - Le propane.....	6
I-1-1 Historique.....	6
I-1-2 Définition.....	6
I-1-3 Composition	7
I-1-4 Utilisation.....	7
I-1-5 Propriétés physiques	8
I-1-6 Combustion du gaz propane	8
I-1-7 Évaporation du gaz propane.....	9
I-2 Le phénomène de tension superficielle	9
I-2-1 Mise en évidence	9
I-2-2 Définition	10
I-2- 3 La tension actif.....	12
I-2- 4 Mesure de la tension superficielle	14
I- 2-4 -1 Méthode de la goutte pendante de la bulle montante.....	14
I-2- 4.2 Méthode de pesée de goutte	14
I-2- 4.3 Méthode de la plaque de Wilhelmy on de l'anneau de du Nouÿ.....	15
II-1 Modélisation	16
II-1-1- Régression linéaire multiple (MLR).....	16
II-2- DEVELOPPEMENT ET EVALUATION DE MODELE.....	18

SOMMAIRE

II-2-1 Sélection d'un sous-ensemble de descripteurs.....	18
II-2-2 Principe.....	18
II-2-3 Initialisation aléatoire du modèle.....	19
II-2-4 Etape de croisement.....	19
II-2-5 Etape de mutation.....	19
II-2-6 Conditions d'arrêt.....	20
II-3 Développement des modèles	20
II-3-1 Paramètres d'évaluation de la qualité de l'ajustement.....	20
II-3-2 Robustesse du modèle.....	21
II-3-3 Domaine d'application.....	21
II-3-4 Test de randomisation.....	22
II-3-5 Validation externe.....	21
III-RESULTATS ET DISCUSSION.....	24
III.1. Calcul du modèle.....	24
III.1.1. Choix du modèle.....	23
II.2. Développement du modèle.....	34
II.2.4. Qualité de l'ajustement.....	35
II.2.3. Diagramme de Williams.....	36
II.2.5. Test de randomisation.....	37
Conclusion générale.....	39
Référence bibliographique.....	40

SOMMAIRE

Liste des tableaux

Tableau I : L'ordres grandeur (cas liq-air).....	12
Tableau III: Valeurs de Tension Superficielle expérimentales, prédites, σ_{ii}, σ_{std}.....	25
Tableau III: Signification des descripteurs.....	35
Tableau V: Les paramètres statistiques pour l'ensemble de calibration (179 composés) et celui de validation (45 composés)	35

SOMMAIRE

LISTE DES FIGURES

<i>Figure01 :force d'attraction des molécules au sein de liquide et a l'interface liquide / aire.</i>	10
<i>Figure 2 : Film de savon formé dans un cadre métallique dont un des côtés est mobile.</i>	11.
<i>Figure 3 : Forces qui s'exercent sur la surface d'une lame de savon</i>	11
<i>Figure 4 : comportement de tensions actives en solution et a l'interface eau l'Aire.....</i>	13
<i>Figure 5 : Evolution de la tension superficielle en fonction de la concentration en tensioactifs en solution.....</i>	14
Figure 06: Qualité de l'ajustement.....	37
Figure 07: Diagramme de Williams.....	37
Figure 08: Test de randomisation.....	38

SYMBOLES ET ABREVIATIONS

Ts	La tension superficielle.
e_i	Résidu : différence entre les valeurs observées et estimées.
$e_{i \text{ std}}$	Résidu de prédiction standardisé.
EQMC	Erreur quadratique moyenne sur l'ensemble de calibration.
EQMP	Erreur quadratique moyenne sur l'ensemble de prédiction.
EQMP _{ext.}	Erreur quadratique moyenne sur l'ensemble de prédiction externe.
F	Statistique de Fisher.
H	Matrice de projection, ou matrice chapeau.
h_{ii}	Éléments diagonaux de la matrice chapeau.
LOO	Cross-validation by leave-one-out: Validation croisée par omission d'une observation.
n-p	Nombre de degrés de liberté de la somme des carrés des résidus.
p	Nombre de descripteurs en comptant la constante (nombre de paramètres).
PRESS	Somme des carrés des erreurs de prédiction.
QSAR	Quantitative Structure/ Activity Relationships.
Q ²	Coefficient de prédiction.
R ²	Coefficient de détermination.
S	Erreur standard.
SCE	Somme des carrés des écarts.
SCT	Somme des carrés totale.

INTRODUCTION GENERALE

Introduction

La tension superficielle est une force qui agit à la surface d'un liquide, à la frontière entre le liquide et l'air ou une autre substance non-miscible. Cette force est créée par les forces d'attraction entre les molécules du liquide, qui attirent les molécules vers leur centre et créent une surface tendue. Cette tension superficielle est mesurée en unité de force par unité de longueur, généralement en newtons par mètre (N/m) ou en dynes par centimètre (dyn/cm). Elle est responsable de plusieurs phénomènes, tels que la formation de gouttes, la capillarité, ou encore l'adhérence ou la cohésion des liquides. La tension superficielle en chimie est une propriété physique des liquides qui est causée par les forces intermoléculaires d'attraction entre les molécules. Elle représente la quantité d'énergie nécessaire pour étendre la surface d'un liquide d'une unité de surface. Elle est difficile, de mesurer la tension superficielle ; et pour meure, il existe différentes méthode dont celle de la goutte montante ou pendante, qui permet une mesure précise de la tension superficielle en fonction du temps et qui donne accès à des informations complémentaire sur la tension actifs.[1]

Dans les derniers années plusieurs méthode à apparaitre, pour la détermination de Ts parmi ces méthode la technique l'approche QSPR (Relations Quantitatives propriété-Structure) est, comme leur nom l'indique, de mettre en place une relation mathématique reliant de manière quantitative des propriétés moléculaires aussi bien électroniques que géométriques, appelées descripteurs, avec une observable macroscopique (activité biologique, toxicité, propriété physico-chimique, etc.), pour une série de composés chimiques similaires à l'aide de méthodes d'analyses de données. Aussi la forme générale de modèle est: L'objectif d'une telle méthode est donc d'analyser les données structurales afin de détecter les facteurs déterminants pour la propriété mesurée. Pour ce faire, différents types d'outils peuvent être employés : régressions multilinéaires (MLR), régressions aux moindres carrés partiels (PLS), arbres de décision, réseaux de neurones, algorithmes génétiques. Une fois cette relation mise en place et validée sur un jeu de validation, elle peut alors être employée pour la prédiction de la propriété de nouvelles molécules, pour lesquelles la valeur

expérimentale n'est pas disponible, voire pour des molécules encore non synthétisées. De tels modèles peuvent également, dans certains cas, être utilisés pour mieux appréhender les phénomènes moléculaires mis en jeu dans la propriété d'intérêt.[2].

Les Avantages de QSPR sont moins coûteux et moins complexes que les expériences physico-chimiques, ce qui peut permettre d'économiser du temps et de l'argent ; On peut fournir des informations moléculaires qui sont difficiles à observer expérimentalement. Les inconvénients de QSPR sont basés sur des hypothèses simplificatrices qui peuvent limiter leur pertinence, on peut être biaisés ou donner des résultats erronés si la base de données utilisée pour leur création n'est pas représentative de la population de molécules souhaitée. Ne Les méthodes QSPR ne peuvent pas capturer toutes les interactions moléculaires, en particulier celles qui sont complexes et difficiles à modéliser.

Notre travail s'inscrit dans ce contexte et on se propose de modéliser la tension superficielle d'un ensemble des dérivées de propane à l'aide de descripteurs moléculaire théoriques en adoptant l'approche hybride algorithmes génétiques/régression linéaire multiple.

Ce travail comporte trois parties

- La première présente des généralités sur les composés étudiés, et la tension superficielle appropriée.
- Dans la deuxième partie nous avons présenté les méthodes de la modélisation moléculaire utilisées lors de l'étude théorique.
- Dans la troisième partie nous avons présenté et discuté le résultat fourni par la modélisation moléculaire.

PARTIE GENERALITES

Généralité

I-1 Le propane

I-1-1 Historique

Le propane est composé de trois atomes de carbone et de huit atomes d'hydrogène (C_3H_8). À l'état naturel, il se trouve dans les dépôts de pétrole associés au gaz naturel qui s'accumule dans les réservoirs souterrains formés de couches poreuses de grès et de roche carbonatée. Le gaz de pétrole liquéfié (GPL) est majoritairement composé de propane et de faibles quantités d'autres hydrocarbures C_3 et C_4 .

On peut utiliser le carburant de propane à de nombreuses fins. À température ambiante, le propane prend la forme d'un gaz. On le transforme facilement en liquide afin de le stocker dans des cylindres de compression fixes ou mobiles. Le propane est un sous-produit des industries du traitement du gaz naturel et du raffinage du pétrole brut. Au Canada, près de 85 à 90 p. 100 du propane provient du traitement du gaz naturel; le reste provient du raffinage du pétrole brut. Le Canada est le quatrième plus grand producteur de propane à l'échelle mondiale : en 2015, il produisait 171,7 millions de barils par jour. Plus de la moitié de ce propane est utilisé au Canada; le reste est exporté aux États-Unis. Le propane est acheminé par gazoduc, par train et par camion ou remorque de transport en vrac. C'est dans les carrefours que les produits sont échangés. Les carrefours canadiens les plus importants se situent à Edmonton et à Sarnia.

I-1-2 Définition

Le propane est un gaz de pétrole liquide, également connu sous le nom de gaz d'hydrocarbure. Comme le butane, il est d'abord extrait sous forme gazeuse. Il est ensuite comprimé en un liquide, sous haute pression. En effet, sous sa forme liquide, le gaz propane est beaucoup plus compact, ce qui le rend plus facile à transporter.

Un gaz peut également devenir liquide en le refroidissant fortement. Le gaz naturel, par exemple, se transforme en une substance liquide à une température de -163 °C. Le gaz propane, quant à lui, se liquéfie à -40 °C. Cependant, la densité plus faible du gaz propane le rend plus facile à comprimer.

Le butane et le propane ont une composition proche. La composition chimique du gaz propane est C_3H_8 et celle du butane est C_4H_{10} . Une différence majeure entre les deux gaz est leur point d'ébullition : le propane a un point d'ébullition beaucoup plus bas. Cela explique pourquoi le propane liquide se transforme également en gaz à basse température.

I-1 -3 Composition

La formule chimique du propane est C_3H_8 , cela signifie qu'une molécule de propane contient 3 atomes de carbone et 8 atomes d'hydrogène. Le propane est une fois et demie plus dense que l'air à température normale et pression atmosphérique, ce qui le différencie du gaz naturel.

Le propane est inodore, mais un additif (l'éthanthiol, ou le méthanthiol) est ajouté afin de permettre une détection aisée des fuites de propane.

Du fait de sa densité supérieure à celle de l'air, une fuite de propane donne lieu à une concentration de propane au sol qui donne naissance à des poches de gaz très inflammables. Une fuite de propane dans un sous-sol peut ainsi remplir la pièce.

I-1-4 Utilisation

Le propane est utilisé dans de nombreux domaines, notamment :

1. **Chauffage** : le propane est un combustible couramment utilisé pour chauffer les maisons, les bâtiments commerciaux et les installations industrielles.
2. **Cuisson** : le propane est également utilisé comme combustible pour les cuisinières et les barbecues.
3. **Énergie alternative** : le propane est une alternative aux énergies fossiles plus polluantes, telles que le charbon ou le pétrole.
4. **Industrie** : le propane est utilisé pour alimenter les équipements de soudure, de découpe et de peinture.
5. **Transport** : les véhicules fonctionnant au propane émettent moins de

polluants que les véhicules fonctionnant à l'essence ou au diesel. Le propane est

également utilisé comme carburant pour les chariots élévateurs et autres équipements de manutention.

Il est important de noter que l'utilisation du propane doit être effectuée dans un environnement bien ventilé et en suivant les précautions de sécurité appropriées pour éviter les fuites, les explosions ou les incendies [3].

I-1-5 Propriétés physiques

Le propane est un gaz incolore et inodore à température ambiante et à pression atmosphérique normale. Il est facilement compressible sous pression atmosphérique et se liquéfie sous forme de gaz liquéfié de pétrole (GPL). Voici quelques propriétés physiques importantes du propane :

- Point d'ébullition : -42,1 °C (-43,8 °F)
- Point de fusion : -187,7 °C (-305,9 °F)
- Densité à l'état liquide : 0,493 g/cm³ à 25 °C
- Densité à l'état gazeux : 1,88 kg/m³ à 25 °C et 1 atm
- Pression de vapeur saturation : 1,7 MPa à 25 °C

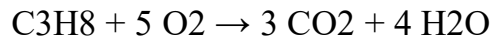
Le propane est également inflammable et peut former des mélanges explosifs avec l'air dans des concentrations spécifiques. Il est donc important de manipuler le propane avec précaution et de suivre les règles de sécurité appropriées lors de sa manipulation et de son utilisation.

I-1-6 Combustion du gaz propane

La combustion du gaz propane, également connue sous le nom de propane, est une réaction chimique exothermique dans laquelle le propane réagit avec l'oxygène

de l'air pour former du dioxyde de carbone, de l'eau et de la chaleur.

L'équation chimique de la combustion du propane est la suivante :



Cette combustion produit de la chaleur, ce qui en fait une source d'énergie courante dans les appareils domestiques tels que les chauffe-eau, les cuisinières et les chauffages. Cependant, il est important de bien ventiler la zone pour éviter toute accumulation de gaz toxiques.

I-1-7 Évaporation du gaz propane

Lorsque le gaz propane est exposé à l'air ou à une source de chaleur, une partie de celui-ci va s'évaporer et se transformer en gaz. L'évaporation se produit en raison de la pression exercée par le gaz dans le réservoir ou la bouteille, ainsi que de la température et de l'humidité de l'environnement.

Lorsqu'il est stocké sous pression, le propane est sous forme liquide dans le réservoir ou la bouteille. L'évaporation se produit lorsque le gaz est libéré de la bouteille ou du réservoir, provoquant une diminution de la température à l'intérieur et forçant le liquide à se vaporiser.

L'évaporation du propane est généralement un processus sûr, mais il est important de prendre des mesures de sécurité pour l'utiliser ou le stocker correctement. Cela inclut l'utilisation d'une bouteille de propane certifiée et d'un régulateur de pression approprié pour contrôler la quantité de gaz qui est libérée. [4]

I-2 Le phénomène de tension superficielle

I-2-1 Mise en évidence

Plusieurs observations simples permettent de mettre en évidence des phénomènes surprenants : un trombone en acier pose à la surface de l'eau ne coule pas alors qu'il a une masse volumique bien supérieure à celle de l'eau. Les gennies, de petits insectes, se déplacent à grande vitesse (-1 m/s) à la surface de l'eau comme s'il s'agissait d'un paroi souple et élastique. Ces observations laissent penser qu'il existe à la surface de l'eau une force qu'il lui

confère des propriétés différentes de celle de son volume.

I-2-2 Définition

Au sein du liquide, une molécule subit de la part de ses voisines des forces d'attraction qui se composent par symétrie (figure 01) à l'interface eau /air la résultante des forces d'attraction.

Interoculaire est cette fois dirigée vers l'intérieure du liquide de la couche superficielle va donc tendance à s'efforce en comprimant le liquide qui réagit pour maintenir à l'équilibre sa surface libre, cette force compensatrice est appelé la tension superficielle et notée γ . Quand le liquide est en contact avec un autre liquide est en contact avec un autre liquide non miscible, on parle de la tension interraciale [5].

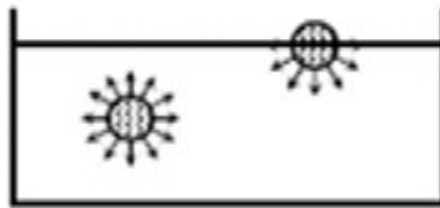


Figure01 :force d'attraction des molécules au sein de liquide et a l'interface liquide / aire.

La non compensation des forces d'attraction à l'interface fait qu'une molécule à la surface est moins stable qu'à l'intérieur du liquide. Le liquide va donc avoir tendance à minimiser sa surface libre pour diminuer son énergie. Pour accroître la surface il faut donc apporter de l'énergie.

Pour quantifier cette énergie à apporter prenons l'exemple d'une lame de savon formée d'un cadre métallique ABCD dont l'un des côtés est mobile sans frottement (Figure 2).

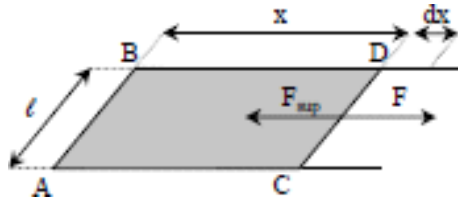


Figure 2 : Film de savon formé dans un cadre métallique dont un des côtés est mobile.

Pour déplacer la portion CD, il faut appliquer une force F pour agrandir la surface du cadre. Cette force s'oppose à la force de tension superficielle qui ramène le cadre à l'équilibre. L'énergie dépensée pour augmenter l'aire superficielle correspond au travail de la force :

$$F = \gamma \cdot l$$

La tension superficielle est donc la force par unité de longueur agissant sur un élément de surface.

Dans de nombreux cas, nous sommes en présence de deux surfaces (lames minces, bulles, ...) (Figure 3) et comme l'expression de la force est valable pour une interface, son expression est :

$$F = 2\gamma \cdot l$$

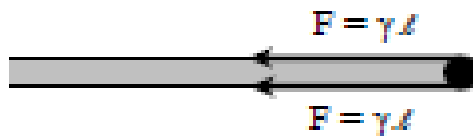


Figure 3 : Forces qui s'exercent sur la surface d'une lame de savon.

Et Comme l'expression de la force est valable comme interface, son expression est la notion de tension superficielle peut aussi être considérée d'un

point de vue énergétique la surface de la lame du schéma, compte-tenu de ses 2 interfaces, vaut $2x.l$ si on veut déplacer le côté Cd d'une longueur dx , il faut fournir un travail

$dW = F \cdot dx$ le travail à fournir par unité de surface créée est donc

$$\frac{dW}{dS} = \frac{F \cdot dx}{2 \cdot l \cdot dx} = \frac{2 \cdot \gamma \cdot l}{2 \cdot l} = \gamma$$

La tension superficielle correspond à l'énergie à fournir pour créer un mètre carré de surface.

Elle s'exprime en $N \cdot m^{-1}$ ce qui équivaut à $J \cdot m^{-2}$ ou $dyne/cm$. L'origine de cette force réside dans les forces intermoléculaires de Van der Waals. Il s'agit de forces dont l'origine est électrostatique qui apparaissent à cause de la polarité de certaines molécules (interaction entre molécules polaires, polarité induite ...). Lorsque la température du liquide augmente la tension interfaciale diminue car les forces de cohésion intermoléculaires sont réduites par l'agitation thermique.

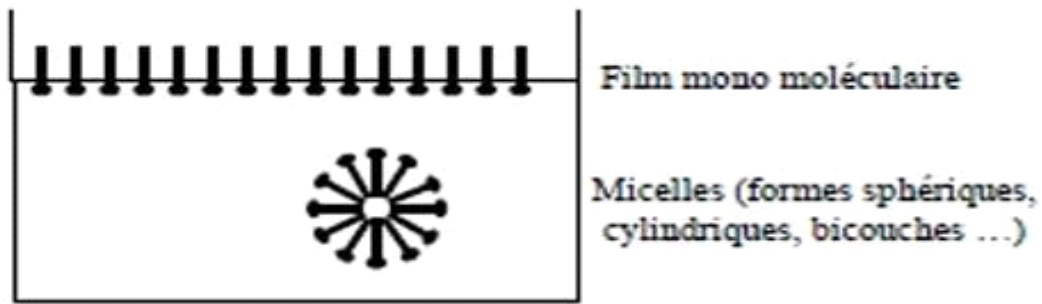
Tableau I : L'ordre de grandeur (cas liq-air)

Liquids	Y(N/m) a 20 C°
Eau	73
Huile végétale	32
Ethanol	22
Ether	17
Mercure	480
L'huile d'olive	

I-2- 3 La tension active

Leur rôle est de diminuer la tension superficielle du liquide dans lequel on les ajoute pour les rendre normalement, moussants, détergents, émulsifiants, ... un tensioactif est composé de 2 parties : une tête hydrophile (groupe polaire ou ionique) attiré par l'eau et une partie hydrophobe (organique) rejeté par l'eau vers la surface.[1]

Il existe quatre grandes catégories de tension active :



Anioniques, cationique, zwitter ioniques et non ioniques en fonction de la nature de leur tête hydrophile. Lorsqu'on interdit une telle molécule dans l'eau, sa partie hydrophobe a tendance à fuir la solution pour se placer spontanément à la surface qui lui procure un espace disponible pour se stabiliser (figure 4)

Figure 4 : comportement de tensions actives en solution et a l'interface eau l'Aire

Quand la concentration en tensioactif augmente dans la phase aqueuse, une partie des molécules tensioactives sont adsorbées à la surface eau/air et la tension superficielle diminue rapidement (Figure 7, zone I).

A partir d'une certaine valeur, il se forme alors un film mono moléculaire de tensioactif à la surface de l'eau et la tension superficielle décroît de façon linéaire avec le logarithme de la concentration (Figure 7, zone II). Lorsque l'interface est saturée, les tensioactifs en solution vont se regrouper en micelle au sein du liquide pour se stabiliser entre eux (Figure 7, zone III). Lorsque la tension superficielle ne diminue plus, on dit qu'on a atteint la concentration micellaire critique (cmc).

La cmc peut être calculée grâce à l'ordonnée de l'intersection des pentes des droites II et III.

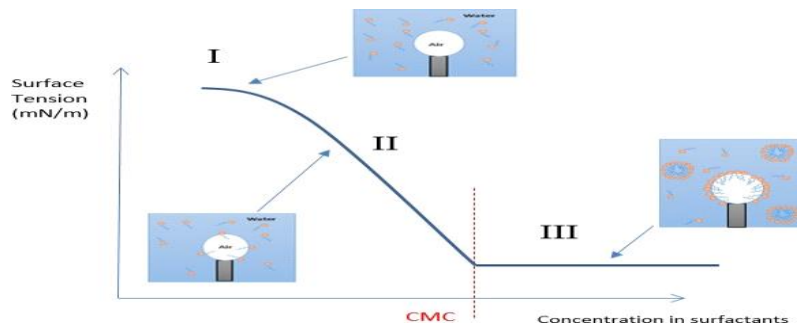


Figure 5 : Evolution de la tension superficielle en fonction de la concentration en tensioactifs en solution

I-2- 4 Mesure de la tension superficielle

Plusieurs techniques peuvent être utilisées pour mesurer la tension superficielle

I- 2-4 -1 Méthode de la goutte pendante de la bulle montante

Une goutte formée ou bout d'un capillaire est observée via une caméra. La forme de la goutte est déterminée par la combinaison de la tension inter faciale et des effets de la gravitation. Les effets de la tension inter faciale forcent la goutte à prendre une forme sphérique tandis que les effets de la gravitation ont tendance à l'allonger pour lui donner une forme en poire dans le cas d'une goutte pendante.

Connaissant la densité du liquide, on en déduit alors la tension superficielle. [6]

I-2- 4.2 Méthode de pesée de goutte

Cette méthode consiste à pas une goutte qui tombe d'un capillaire de rayon r connu. En premier approximation les forces qui s'appliquent sur la goutte sont son poids $P=mg$ et la force due à la tension superficielle Y au niveau du capillaire

Au moment précis au la goutte se détache, le poids de la goutte est égal aux forces capillaire $P=F$. On retrouve la loi de tâte

$$\gamma = \frac{mg}{2\pi r}$$

$$F=2\pi r\gamma.$$

I-2- 4.3 Méthode de la plaque de Wilhelmy ou de l'anneau de du Nouÿ

On tire hors du liquide une plaque ou un anneau tout en mesurant la force exercée. On note la valeur de la force juste avant que le ménisque se détache. On divise la force ainsi obtenue par la largeur de la plaque ou par le périmètre de l'anneau, et on obtient la valeur de la tension superficielle module un facteur trigonométrique.

Chaque technique a ses avantages et permet de mesurer la tension superficielle dans une certaine gamme de valeurs. Cependant, la méthode de la goutte pendante ou montante est beaucoup plus précise et apporte en plus des informations sur l'absorption des tensioactifs à la surface.

Le suivi de l'évolution de la forme la goutte par caméra permet d'avoir la dynamique d'évolution de la tension superficielle et de remonter aux paramètres propres aux tension actifs : le volume occupé par une tête hydrophile, l'air inter faciale développée par une molécule tensioactive, le nombre de molécules adsorbées à l'interface... volume occupé par une tête hydrophile, l'air inter faciale développé par une molécule tensioactive, le nombre de molécules adsorbées à l'interface peuvent

MODELISATION

II-1 Modélisation

La modélisation est un outil destiné aux chercheurs préoccupés par la structure et la réactivité des molécules. La connaissance de la structure des édifices moléculaires permet de comprendre ce qui est réalisé dans une transformation physique, chimique ou biologique. Elle peut permettre aussi de prévoir de telles transformations. La compréhension comme la prévision sont considérablement facilitées lorsque l'on peut visualiser les structures. Une molécule est correctement décrite par sa géométrie et ses propriétés thermodynamiques. La visualisation doit rendre compte de l'ensemble de ces caractéristiques.

La question essentielle est de représenter une molécule sur l'écran de la façon la plus proche possible de la "réalité".

L'utilisation de l'informatique a permis de mettre au point un outil performant : la modélisation moléculaire.

Le but de la régression simple (resp. multiple) est d'expliquer une variable Y à l'aide d'une variable X (resp. plusieurs variables X1, ..., Xq). La variable Y est appelée variable dépendante, ou variable à expliquer et les variables Xj (j=1,...,q) sont appelées variables indépendantes, ou variables explicatives.

II-1-1- Régression linéaire multiple (MLR) [6-7]

La régression linéaire multiple est la méthode la plus simple de modélisation. elle consiste à rechercher une équation linéaire par rapport à ses paramètres reliant la variable à modéliser au vecteur d'entrées $x = \{ x_k, k = 1, \dots, q \}$. Ces entrées peuvent être des fonctions non paramétrées, ou à paramètres fixés, de ces variables. L'équation linéaire recherchée est de la forme :

$$g(x, \theta) = \sum_{k=1}^q \theta_k x_k = X\theta \quad (1)$$

où $\theta = \{\theta_k, k=1, \dots, q\}$ est le vecteur des paramètres; X , matrice des observations de taille (N, q) , est définie comme la matrice dont les éléments de la colonne k prennent pour valeurs les N mesures de la variable k . Pour chaque élément i de la base d'apprentissage, le résidu est défini comme la différence entre la valeur de la grandeur à modéliser pour cet élément i et l'estimation du modèle :

$$R_i = y^i - g(x^i, \theta) \quad (2)$$

L'apprentissage est réalisé par minimisation de la fonction de coût des moindres carrés, qui mesure l'ajustement du modèle g aux données d'apprentissage :

$$J(\theta) = \sum_{i=1}^N (R_i)^2 = \sum_{i=1}^N [y^i - g(x^i, \theta)]^2 = \|y - X\theta\|^2 \quad (3)$$

La fonction $J(\theta)$ est une fonction positive quadratique en θ : son minimum est unique. Il est donné par :

$$\theta_{mc} = (X^T X)^{-1} X^T y \quad (4)$$

Les paramètres θ_k sont appelés coefficients de régression partielle ; chacun d'eux mesure l'effet de la variable explicative x_k concernée sur la propriété modélisée lorsque les autres variables explicatives sont maintenues constantes.

La régression linéaire est facile à mettre en œuvre, et les coefficients θ_k obtenus peuvent être interprétés : ils mesurent l'influence de chacune des variables sur les grandeur étudiée.

II-2- DEVELOPPEMENT ET EVALUATION DE MODELE.**II-2-1 Sélection d'un sous-ensemble de descripteurs.**

Des logiciels spécialisés permettent le calcul de plus de 6000 descripteurs moléculaires appartenant à différentes classes. Plutôt que de chercher à expliquer la variable dépendante (grandeur d'intérêt) par tous les régresseurs (descripteurs moléculaires), on peut chercher seulement un ensemble réduit de régresseurs qui donne une reconstitution aussi satisfaisante de la variable à expliquer. Parmi les stratégies mises en œuvre pour la sélection d'un ensemble réduit de variables explicatives, on peut citer : les méthodes de pas-à-pas (méthode descendante; méthode ascendante, et méthode dite stepwise), ainsi que les algorithmes génétiques.

En général, la comparaison se fait à l'avantage des algorithmes génétiques (AG) que nous avons appliqués dans le présent travail, et que nous rappelons succinctement.

II-2-2 Principe.

Dans la terminologie des algorithmes génétiques, le vecteur binaire I , appelé "chromosome", est un vecteur de dimension p où chaque position (un "gène") correspond à une valeur (1 si elle figure dans le modèle, 0 sinon). Chaque chromosome représente un modèle basé sur un ensemble de variables explicatives.

On commence par définir le paramètre statistique à optimiser (par exemple maximiser Q^2 en utilisant la validation croisée par "leave-one-out" ; (cf. infra), avec la taille P de la population du modèle (par exemple, $P = 100$), et le nombre maximum de variables L permises pour le modèle (par exemple, $L = 10$) ; le minimum de variables permises est généralement supposé égal à 1. De plus, une probabilité de croisement p_c (habituellement élevée $p_c = 0,9$), et une probabilité de mutation p_M (habituellement faible, $p_M = 0,1$) doivent être également définies.

Après définition des principaux paramètres, la mise en œuvre de l'algorithme génétique est démarrée, son évolution comprend trois étapes principales.

II-2-3 Initialisation aléatoire du modèle.

La population est constituée au départ de modèles aléatoires avec des variables comprises entre 1 et L, puis les modèles sont ordonnés eu égard au paramètre statistique sélectionné – la qualité du modèle – (le meilleur modèle est en première position, le plus mauvais en position P) ;

II-2-4 Etape de croisement.

A partir de la population, on sélectionne des paires de modèles (aléatoirement, ou avec une probabilité proportionnelle à leur qualité). Puis, pour chaque paire de modèles on conserve les caractéristiques communes, c'est-à-dire les variables exclues dans les 2 modèles restent exclues, et les variables intégrées dans les 2 modèles sont conservées. Pour les variables sélectionnées dans un modèle et éliminées dans l'autre, on en essaye un certain nombre au hasard que l'on compare à la probabilité de croisement p_c : si le nombre aléatoire est inférieur à la probabilité de croisement, la variable exclue est intégrée au modèle et vice versa. Finalement, le paramètre statistique du nouveau modèle est calculé : si la valeur de ce paramètre est meilleure que la plus mauvaise pour la population, le modèle est intégré à la population, à la place correspondant à son rang ; dans la cas contraire, il n'est pas pris en compte. Cette procédure est répétée pour de nombreuses paires (100 fois par exemple).

II-2-5 Etape de mutation.

Pour chaque modèle de la population (c'est-à-dire pour chaque chromosome) p nombres aléatoires sont éprouvés, et, un à la fois, chacun est comparé à la probabilité de mutation, p_M , définie : chaque gène demeure inchangé si le nombre aléatoire correspondant excède la probabilité de mutation, dans le cas contraire, on le change de 0 à 1 ou vice versa. Les faibles valeurs de p_M permettent uniquement peu de mutations, conduisant à de nouveaux chromosomes peu différents des chromosomes générateurs.

Après obtention du modèle transformé, on en calcule le paramètre statistique : si cette valeur est meilleure que la plus mauvaise de la population, le modèle est intégré à la population, à la place correspondant à son rang ; dans le cas contraire il n'est pas pris en compte.

II-2-6 Conditions d'arrêt.

Les étapes de croisement et de mutation sont répétées jusqu'à la rencontre d'une condition d'arrêt (par exemple un nombre maximum d'itérations défini par l'utilisateur), ou qu'il est mis fin arbitrairement au processus.

Une caractéristique importante de la sélection d'un ensemble réduit de variables par algorithme génétique est qu'on n'obtient pas nécessairement un modèle unique, mais le résultat consiste habituellement en une population de modèles acceptables ; cette caractéristique, parfois considérée comme un désavantage, fournit une opportunité pour procéder à une évaluation des relations avec la réponse à différents points de vue.

II-3 Développement des modèles

Les techniques les plus courantes pour établir des modèles QSPR utilisent l'analyse de régression (régression linéaire multiple : RLM ; projection des structures latentes par les moindres carrés partiels : PLS), les réseaux neuronaux (RNA), et machine à support de vecteur (SVM).

II-3-1 Paramètres d'évaluation de la qualité de l'ajustement.

Deux paramètres sont couramment utilisés :

Le coefficient de détermination multiple :

$$R^2 = 1 - \frac{SCE}{SCT} = 1 - \frac{\sum_1^n (y_i - \hat{y}_i)^2}{\sum_1^n (y_i - \bar{y})^2} \quad (6)$$

où \hat{y}_i est la valeur estimée du paramètre physique, et \bar{y} la moyenne des valeurs observées.

La racine de l'erreur quadratique moyenne de prédiction (désignée également par EQMP) :

$$EQMP = \sqrt{\frac{1}{n} \sum_1^n (y_i - \hat{y}_{(i)})^2} \quad (7)$$

II-3-2 Robustesse du modèle.

La stabilité du modèle a été explorée en utilisant la "validation croisée par omission d'une observation" (LOO : cross-validation by leave-one-out) [10]. Elle consiste à recalculer le modèle sur $(n - 1)$ composés de calibrage, le modèle obtenu servant alors à estimer la valeur de la propriété du composé éliminé noté $\hat{y}_{(i)}$. On répète le procédé pour chacun des n composés de l'ensemble de calibrage.

"La somme des carrés des erreurs de prédiction", désignée par l'acronyme PRESS (pour : Predictive Residual Sum of Squares) :

$$\text{PRESS} = \sum_{i=1}^n (y_i - \hat{y}_{(i)})^2 \quad (8)$$

est une mesure de la dispersion de ces estimations. On l'utilise pour définir le coefficient de prédiction :

$$Q_{\text{LOO}}^2 = \frac{\text{SCT} - \text{PRESS}}{\text{SCT}} \quad (9)$$

Contrairement à R^2 , qui augmente avec le nombre de paramètres de la régression, le facteur Q_{LOO}^2 affiche une courbe avec maximum (ou avec palier), obtenu pour un certain nombre de variables explicatives, puis décroît par la suite de façon monotone. Ce fait confère une grande importance au coefficient Q_{LOO}^2 . Une valeur de $Q_{\text{LOO}}^2 > 0,5$ est, généralement, considérée comme satisfaisante, et une valeur supérieure à 0,9 est excellente [11].

II-3-3 Domaine d'application.

Le domaine d'application a été discuté à l'aide du diagramme de Williams (traité en détail dans [12, 13], représentant les résidus de prédiction standardisés en fonction des valeurs des leviers h_i . L'équation (10) définit le levier d'un composé dans l'espace original des variables indépendantes (x_i)

$$h_i = \mathbf{x}_i (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_i^T \quad (10)$$

Où (x_i) est le vecteur ligne des descripteurs du composé i et \mathbf{X} ($n \times p$) la matrice du modèle déduite des valeurs des descripteurs de l'ensemble de calibrage ; l'indice T désigne le vecteur (ou la matrice) transposé (e).

La valeur critique du levier (h^*) est fixée à $3(p+1)/n$. Si $h_{ii} < h^*$, la probabilité d'accord entre les valeurs mesurée et prédite du composé i est aussi élevée que celle des composés de calibrage. Les composés avec $h_{ii} > h^*$ renforcent le modèle quand ils appartiennent à l'ensemble de calibrage, mais auront, sinon, des valeurs prédites douteuses sans pour autant être forcément aberrantes, les résidus pouvant être bas.

II-3-4 Test de randomisation

Ce test permet de mettre en évidence des corrélations dues au hasard. Il consiste à générer un vecteur « propriété considérée » par permutation aléatoire des composantes du vecteur réel. On calcule alors sur le vecteur obtenu (considéré comme vecteur expérimental réel) un modèle QSAR, selon la méthode habituelle. Ce procédé est répété plusieurs fois (100 dans notre cas) [14].

II-3-5 Validation externe.

En plus du test de randomisation, il est intéressant [15], pour juger de la qualité du modèle, de considérer le coefficient de prédiction externe calculé comme suit :

$$Q^2_{ext} = 1 - \frac{\sum_{i=1}^{n_{ext}} (\hat{y}_{i/i} - y_i)^2 / n_{ext}}{\sum_{i=1}^{n_{tr}} (y_i - \bar{y}_{tr})^2 / n_{tr}} \quad (11)$$

la racine de l'écart quadratique moyen (RMSE pour Root Mean Squared Error), calculée sur différents ensembles:

- Ensemble d'estimation (appelée EQMC)
- Ensemble de prédiction externe (désignée par EQMPext).

Ces valeurs RMSE sont mieux adaptées, pour juger de la qualité d'un modèle que les valeurs de R^2 et Q^2 seules, qui constituent de bons tests uniquement pour des données réparties régulièrement.

$$EQMC = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (12)$$

$$EQMP_{ext} = \sqrt{\frac{\sum_{i=1}^{n_{ext}} (y_i - \hat{y}_i)^2}{n_{ext}}} \quad (13)$$

RESULTATS ET DISCUSSIONS

Dans cette partie, nous allons déterminer la tension superficielle d'une série de dérive de propane par modélisation d'un ensemble d'estimation. L'ensemble des données ont été retiré par l'article [16]. Le meilleur modèle obtenue sera validé par l'approche hybride algorithme génétique et régression linéaire AG/RLM.

La liste des molécules étudiées dans notre travail se présente dans le Tableau II, cette base des données contient 224 composés, avec la propriété à étudier et les variables explicatives.

III-1 Calcul du modèle

II-3-1 Choix du modèle

Les molécules sont dessinées par le logiciel ChemDraw puis elles sont optimisées en utilisant le logiciel HyperChem [17]. Les descripteurs moléculaires ont été calculés à l'aide du logiciel informatique DRAGON [18], plus de 1600 descripteurs sont calculés.

L'ensemble des données a été décomposé en deux sous-ensembles selon article, 179 composés pour la construction du modèle et 45 composés pour la validation externe.

D'après l'algorithme génétique dans la version MOBYDIGS [19], plusieurs modèles sont obtenues pour chaque jeu de groupe; le choix a été opté pour le modèle qui conduit aux meilleurs statistiques des 50 modèles générés par algorithmes génétiques.

Les valeurs du point de fusion expérimentales et prédites, les valeurs des Leviers, et les valeurs des erreurs standardisées sont présentées dans le tableau II au-dessous:

Tableau II: Valeurs de Tension Superficielle expérimentales, prédites,h_{ii}, e_{istd}

N	Composé	Status	Y Exp	Y-Calcul	Y-Pred	Hat	Err. Calc	Err. Pred	Std.Err. Calc	Std.Err. r. Pred
01	1,1,1,2,2,3,3,4,4-Nonafluorobutane	Training	10.50	9.83	9.81	0.022	-0.67	-0.69	-0.53	-0.55
02	1,1,1,2,2,3,3,3Heptachloro-3-fluoropropane	Training	41.80	41.56	41.54	0.100	-0.24	-0.26	-0.20	-0.22
03	1,1,1,2,2,3-Hexachloro-3-fluoropropane	Training	38.40	37.44	37.38	0.066	-0.96	-1.02	-0.78	-0.83
04	1,1,1,2,2,3-Hexafluoropropane	Training	0.30	10.33	10.33	0.015	0.03	0.03	0.03	0.03
05	1,1,1,2,3,3,3-Heptafluoropropane	Training	9.60	9.19	9.18	0.020	-0.41	-0.42	-0.32	-0.33
06	1,1,1,2,3,3-Hexafluoropropane	Training	10.00	10.19	10.19	0.015	0.19	0.19	0.15	0.15
07	1,1,1,2-Tetrachloro-2,2-difluoroethane	Training	29.20	30.36	30.39	0.027	1.16	1.19	0.93	0.95
08	1,1,1,2-Tetrachloro-2-fluoroethane	Training	31.40	31.36	31.36	0.026	-0.04	-0.04	-0.03	-0.03
09	1,1,1,2-Tetrachloroethane	Training	34.90	32.32	32.13	0.069	-2.58	-2.77	-2.10	-2.26
10	1,1,1,2-Tetrafluoroethane	Training	9.10	9.61	9.63	0.053	0.51	0.53	0.41	0.43
11	1,1,1,3,3-Pentachloro-2,2,3-trifluoropropane	Training	31.00	32.24	32.29	0.041	1.24	1.29	0.99	1.03
12	1,1,1,3,3-Pentachloro-2,2-difluoropropane	Training	32.80	34.21	34.28	0.050	1.41	1.48	1.13	1.19
13	1,1,1,3,3-Pentafluorobutane	Training	12.40	11.82	11.82	0.011	-0.58	-0.58	-0.46	-0.46
14	1,1,1,3,3-Pentafluoropropane	Training	10.70	11.36	11.38	0.032	0.66	0.68	0.53	0.54
15	1,1,1,3-Tetrachloro-2,2-difluoropropane	Training	29.80	27.22	27.15	0.028	-2.58	-2.65	-2.05	-2.11
16	1,1,1-Trichloro2,2,3,3,3-pentafluoropropane	Training	21.20	23.03	23.06	0.016	1.83	1.86	1.45	1.47
17	1,1,1-Trichloro-2,2,3,3-tetrafluoropropane	Training	22.30	23.71	23.73	0.014	1.41	1.43	1.12	1.13
18	1,1,1-Trichloro-2,2,3-	Training	24.20	21.73	21.71	0.009	-2.47	-2.49	-1.95	-1.97

	trifluoropropane									
19	1,1,1-Trichloro-2-fluoroethane	Training	27.40	26.53	26.52	0.015	-0.87	-0.88	-0.69	-0.70
20	1,1,1-Trichloroethane	Training	28.90	25.36	25.30	0.017	-3.54	-3.60	-2.80	-2.85
21	1,1,1Trichlorotrifluoroethane	Training	23.00	26.09	26.22	0.038	3.09	3.22	2.48	2.58
22	1,1,1-Trifluoroethane	Training	9.00	10.28	10.30	0.014	1.28	1.30	1.01	1.03
23	1,1,2,2,3-Pentafluoropropane	Training	10.70	7.59	7.45	0.045	-3.11	-3.25	-2.50	-2.61
24	1,1,2,2-Tetrachloro-1,2-difluoroethane	Training	29.20	30.82	30.86	0.024	1.62	1.66	1.29	1.32
25	1,1,2,2-Tetrachloro-1-fluoroethane	Training	31.40	31.80	31.81	0.026	0.40	0.41	0.32	0.33
26	1,1,2,2-Tetrachloroethane	Training	33.90	34.55	34.58	0.055	0.65	0.68	0.52	0.55
27	1,1,2,2Tetrafluoro propane	Training	10.80	11.84	11.85	0.013	1.04	1.05	0.82	0.83
28	1,1,2,3,3Pentafluoro propane	Training	10.42	10.71	10.71	0.013	0.29	0.29	0.23	0.23
29	1,1,2-Trichloro-1,2-difluoroethane	Training	24.60	24.76	24.76	0.013	0.16	0.16	0.13	0.13
30	1,1,2-Trichloro-1-fluoroethane	Training	27.40	25.76	25.73	0.014	-1.64	-1.67	-1.30	-1.32
31	1,1,2-Trichloro-2-fluoroethane	Training	26.50	25.66	25.65	0.014	-0.84	-0.85	-0.66	-0.67
32	1,1,2-Trichloroethane	Training	29.77	26.54	26.48	0.018	-3.23	-3.29	-2.56	-2.61
33	1,1,2Trichlorotrifluoroethane	Training	23.00	24.95	24.98	0.015	1.95	1.98	1.54	1.57
34	1,1,2-Trifluoroethane	Training	9.50	9.09	9.08	0.019	-0.41	-0.42	-0.33	-0.33
35	1,1,3,3-Tetrachloro-2,2-difluoropropane	Training	29.00	28.67	28.66	0.024	-0.33	-0.34	-0.26	-0.27
36	1,1,3-Trichloro-1,2,2,3-tetrafluoropropane	Training	22.30	23.07	23.08	0.016	0.77	0.78	0.61	0.62
37	1,1,3-Trichloro-1,2,2-trifluoropropane	Training	24.20	24.16	24.16	0.014	-0.04	-0.04	-0.03	-0.03
38	1,1,3-Trichloro-2,2,3-trifluoropropane	Training	23.60	24.04	24.05	0.013	0.44	0.45	0.35	0.36
39	1,1-Dichloro-1,2,2,3-tetrafluoropropane	Training	19.10	19.10	19.10	0.010	0.00	0.00	0.00	0.00

40	1,1-Dichloro-1,2,2-trifluoroethane	Training	18.40	18.81	18.82	0.010	0.41	0.42	0.33	0.33
41	1,1-Dichloro-1,2,2-trifluoropropane	Training	19.70	20.18	20.19	0.009	0.48	0.49	0.38	0.39
42	1,1-Dichloro-1,2-difluoroethane	Training	20.50	19.39	19.38	0.009	-1.11	-1.12	-0.88	-0.89
43	1,1-Dichloro-1,2-difluoropropane	Training	20.90	20.63	20.63	0.011	-0.27	-0.27	-0.21	-0.22
44	1,1-Dichloro-1,3,3,3-tetrafluoropropane	Training	19.10	19.07	19.07	0.010	-0.03	-0.03	-0.03	-0.03
45	1,1-Dichloro-1-fluoroethane	Training	21.40	20.46	20.45	0.010	-0.94	-0.95	-0.74	-0.75
46	1,1-Dichloro-2,2,3-trifluoropropane	Training	20.20	20.08	20.08	0.009	-0.12	-0.12	-0.09	-0.09
47	1,1-Dichloro-2,2-difluoroethane	Training	19.70	19.20	19.19	0.010	-0.50	-0.51	-0.40	-0.40
48	1,1-Dichloro-2,2-difluoroethylene	Training	20.00	23.65	23.72	0.019	3.65	3.72	2.89	2.95
49	1,1-Dichloro-2,2-difluoropropane	Training	20.90	20.67	20.66	0.011	-0.23	-0.24	-0.18	-0.19
50	1,1-Dichloroethane	Training	23.20	21.36	21.33	0.015	-1.84	-1.87	-1.46	-1.48
51	1,1-Dichlorotetrafluoro ethane	Training	17.30	17.75	17.75	0.014	0.45	0.45	0.35	0.36
52	1,1-Difluoroethane	Training	9.50	10.23	10.24	0.019	0.73	0.74	0.58	0.59
53	1,1-Difluoroethylene	Training	8.70	11.36	11.40	0.017	2.66	2.70	2.10	2.14
54	1,2,2,3-Tetrachloro-1,1,3,3-tetrafluoropropane	Training	25.90	27.09	27.13	0.031	1.19	1.23	0.95	0.98
55	1,2-Dibromo-1,1-difluoroethane	Training	29.30	27.36	27.33	0.015	-1.94	-1.97	-1.53	-1.56
56	1,2-Dibromo-1-fluoroethane	Training	31.90	27.97	27.90	0.017	-3.93	-4.00	-3.12	-3.17
57	1,2-Dibromotetrafluoro ethane	Training	24.40	25.79	25.82	0.020	1.39	1.42	1.10	1.12
58	1,2-Dichloro-1,1,2,3,3-pentafluoropropane	Training	17.60	18.46	18.47	0.013	0.86	0.87	0.68	0.69
59	1,2-Dichloro-1,1,2,3-tetrafluoropropane	Training	19.10	19.02	19.02	0.009	-0.08	-0.08	-0.07	-0.07
60	1,2-Dichloro-1,1,2-trifluoroethane	Training	18.40	18.75	18.75	0.010	0.35	0.35	0.28	0.28

61	1,2-Dichloro-1,1,3,3,3-pentafluoropropane	Training	17.60	18.42	18.43	0.013	0.82	0.83	0.65	0.65
62	1,2-Dichloro-1,1-difluoropropane	Training	20.90	20.54	20.53	0.010	-0.36	-0.37	-0.29	-0.29
63	1,2-Dichloro-1,2,3,3-tetrafluoropropane	Training	18.50	18.91	18.91	0.009	0.41	0.41	0.32	0.32
64	1,2-Dichloro-1-fluoroethane	Training	22.10	20.06	20.04	0.013	-2.04	-2.06	-1.61	-1.63
65	1,2-Dichloro-2-fluoropropane	Training	23.00	21.73	21.71	0.015	-1.27	-1.29	-1.01	-1.02
66	1,2-Dichlorotetrafluoro ethane	Training	17.30	17.75	17.76	0.014	0.45	0.46	0.36	0.36
67	1,2-Difluoroethane	Training	10.60	10.18	10.18	0.019	-0.42	-0.42	-0.33	-0.34
68	1,3,3-Trichloro-1,1,2,2-tetrafluoropropane	Training	22.30	23.04	23.05	0.016	0.74	0.75	0.58	0.59
69	1,3-Dichloro-1,1,2,2,3,3-hexafluoropropane	Training	16.80	17.46	17.47	0.019	0.66	0.67	0.52	0.53
70	1,3-Dichloro-1,1,2,2-tetrafluoropropane	Training	19.10	19.02	19.02	0.009	-0.08	-0.08	-0.07	-0.07
71	1,3-Dichloro-1,1,2,3,3-pentafluoropropane	Training	17.60	18.48	18.49	0.013	0.88	0.89	0.69	0.70
72	1,3-Dichloro-1,1,3,3-tetrafluoropropane	Training	19.12	19.05	19.05	0.010	-0.07	-0.07	-0.05	-0.05
73	1,3-Dichloro-1,2,2,3-tetrafluoropropane	Training	18.50	18.89	18.90	0.009	0.39	0.40	0.31	0.31
74	1-Chloro-1,1,2,2,3,3,3-heptafluoropropane	Training	12.80	12.97	12.97	0.021	0.17	0.17	0.13	0.13
75	1-Chloro-1,1,2,2,3,3-hexafluoropropane	Training	13.40	13.48	13.48	0.015	0.08	0.08	0.06	0.06
76	1-Chloro-1,1,2,2,3-pentafluoropropane	Training	14.40	14.57	14.57	0.010	0.17	0.17	0.14	0.14
77	1-Chloro-1,1,2,2-tetrafluoroethane	Training	13.00	12.53	12.52	0.014	-0.47	-0.48	-0.37	-0.38
78	1-Chloro-1,1,2,3,3,3-hexafluoropropane	Training	13.40	13.43	13.43	0.015	0.03	0.03	0.02	0.02
79	1-Chloro-1,1,2,3-tetrafluoropropane	Training	15.20	15.03	15.02	0.009	-0.17	-0.18	-0.14	-0.14
80	1-Chloro-1,1,2-	Training	14.30	13.69	13.69	0.011	-0.61	-0.61	-0.48	-0.49

	Trifluoroethane									
81	1-Chloro-1,1,2-trifluoropropane	Training	15.50	16.08	16.09	0.010	0.58	0.59	0.46	0.46
82	1-Chloro-1,1,3,3-tetrafluoropropane	Training	15.10	15.06	15.06	0.009	-0.04	-0.04	-0.04	-0.04
83	1-Chloro-1,1,3-trifluoropropane	Training	16.58	16.13	16.13	0.010	-0.45	-0.45	-0.35	-0.35
84	1-Chloro-1,1-difluoroethane	Training	14.70	14.79	14.79	0.011	0.09	0.09	0.07	0.07
85	1-Chloro-1,2,2,3,3-pentafluoropropane	Training	14.00	14.42	14.42	0.010	0.42	0.42	0.33	0.33
86	1-Chloro-1,2,2-Trifluoroethane	Training	13.70	13.58	13.57	0.012	-0.12	-0.13	-0.10	-0.10
87	1-Chloro-1,2,2-trifluoropropane	Training	15.50	16.05	16.05	0.010	0.55	0.55	0.43	0.44
88	1-Chloro-1,2-difluoroethane	Training	15.30	14.58	14.57	0.012	-0.72	-0.73	-0.57	-0.58
89	1-Chloro-1-fluoroethane	Training	15.80	15.17	15.16	0.014	-0.63	-0.64	-0.50	-0.50
90	1-Chloro-2,2,2-Trifluoroethane	Training	14.30	13.48	13.47	0.012	-0.82	-0.83	-0.65	-0.65
91	1-Chloro-2,2,3-trifluoropropane	Training	16.50	16.23	16.23	0.010	-0.27	-0.27	-0.21	-0.22
92	1-Chloro-2,2-difluoropropane	Training	17.00	17.30	17.30	0.014	0.30	0.30	0.24	0.24
93	2,2,2-Trifluoroethyl methyl ether	Training	13.90	14.35	14.37	0.032	0.45	0.47	0.36	0.37
94	2,2-Dichloro-1,1,1,3-tetrafluoropropane	Training	19.10	19.03	19.03	0.010	-0.07	-0.07	-0.05	-0.05
95	2,2-Dichloro-1,1,1-trifluoroethane	Training	18.40	18.69	18.69	0.011	0.29	0.29	0.23	0.23
96	2,2-Dichloro-1,1,3,3-tetrafluoropropane	Training	18.50	18.94	18.95	0.009	0.44	0.45	0.35	0.35
97	2,3,3,3-Tetrafluoropropene	Training	10.30	12.49	12.53	0.018	2.19	2.23	1.74	1.77
98	2,3-Dichloro-1,1,1,3-tetrafluoropropane	Training	18.50	18.83	18.83	0.009	0.33	0.33	0.26	0.26
99	2,3-Dichloro-1,1,1-trifluoropropane	Training	20.20	20.05	20.05	0.009	-0.15	-0.15	-0.12	-0.12
100	2-Chloro-1,1,1,2,3,3,3-	Training	12.80	12.94	12.94	0.021	0.14	0.14	0.11	0.11

	heptafluoropropane									
101	2-Chloro-1,1,1,2-tetrafluoroethane	Training	13.00	12.46	12.46	0.015	-0.54	-0.54	-0.42	-0.43
102	2-Chloro-1,1,1,3,3-pentafluoropropane	Training	14.00	14.42	14.42	0.010	0.42	0.42	0.33	0.33
103	2-Chloro-1,1,3,3-tetrafluoropropane	Training	14.70	14.94	14.94	0.009	0.24	0.24	0.19	0.19
104	2-Chloro-1fluoropropane	Training	18.20	17.52	17.50	0.020	-0.68	-0.70	-0.54	-0.55
105	2-Chloro-2-fluoropropane	Training	17.60	17.58	17.58	0.021	-0.02	-0.02	-0.01	-0.01
106	3-Chloro-1,1,1,2,2,3-hexafluoropropane	Training	13.40	13.42	13.42	0.015	0.02	0.02	0.01	0.01
107	3-Chloro-1,1,1,2,3-pentafluoropropane	Training	14.00	14.37	14.37	0.010	0.37	0.37	0.29	0.29
108	3-Chloro-1,1,1,3-tetrafluoropropane	Training	15.10	15.00	15.00	0.009	-0.10	-0.10	-0.08	-0.08
109	3-Chloro-1,1,1-trifluoropropane	Training	16.50	16.15	16.15	0.010	-0.35	-0.35	-0.28	-0.28
110	3-Chloro-1,1,2,3-tetrafluoropropane	Training	14.73	14.88	14.88	0.009	0.15	0.15	0.12	0.12
111	3-Chloro-1,1-difluoropropane	Training	17.63	17.19	17.18	0.013	-0.44	-0.45	-0.35	-0.35
112	3-Chloro-1,3,3-trifluoropropane	Training	16.58	16.13	16.13	0.010	-0.45	-0.45	-0.35	-0.35
113	Bis(difluoromethyl)ether	Training	10.80	11.09	11.12	0.095	0.29	0.32	0.24	0.26
114	Bromodifluoromethane	Training	16.60	15.53	15.45	0.063	-1.07	-1.15	-0.87	-0.93
115	Bromotrifluoromethane	Training	15.30	13.94	13.83	0.075	-1.36	-1.47	-1.11	-1.20
116	Butane	Training	16.75	17.46	17.50	0.065	0.71	0.75	0.57	0.61
117	Carbontetrachloride (Tetrachloromethane)	Training	35.20	36.70	36.91	0.121	1.50	1.71	1.26	1.43
118	Chlorodifluoromethane	Training	12.20	11.72	11.70	0.046	-0.48	-0.50	-0.38	-0.40
119	Chloroethane	Training	17.90	16.15	16.11	0.020	-1.75	-1.79	-1.39	-1.42
120	Chloroethylene	Training	17.50	19.63	19.67	0.019	2.13	2.17	1.69	1.72
121	Chlorofluoroethane	Training	15.80	15.12	15.11	0.015	-0.68	-0.69	-0.54	-0.54
122	Chlorofluoropropane	Training	18.20	17.76	17.75	0.019	-0.44	-0.45	-0.35	-0.35
123	Chloroheptafluorocyclobutane	Training	14.98	14.59	14.58	0.032	-0.39	-0.40	-0.31	-0.32

124	Chlorohexafluoropropane	Training	13.40	13.42	13.42	0.015	0.02	0.02	0.01	0.01
125	Chloromethane	Training	14.64	15.93	15.97	0.028	1.29	1.33	1.03	1.06
126	Chloropentafluoropropane	Training	14.40	14.59	14.59	0.010	0.19	0.19	0.15	0.15
127	Chlorotetrafluoropropane	Training	14.70	15.04	15.04	0.009	0.34	0.34	0.26	0.27
128	Chlorotrifluoroethylene	Training	13.50	15.77	15.81	0.019	2.27	2.31	1.80	1.83
129	Chlorotrifluoromethane	Training	11.40	12.60	12.63	0.023	1.20	1.23	0.95	0.98
130	Decafluorobutane	Training	10.20	8.77	8.72	0.030	-1.43	-1.48	-1.14	-1.18
131	Dibromodifluoromethane	Training	28.50	28.21	28.18	0.086	-0.29	-0.32	-0.24	-0.26
132	Dichlorodifluoroethane	Training	20.50	19.20	19.19	0.010	-1.30	-1.31	-1.02	-1.03
133	Dichlorodifluoromethane	Training	18.30	18.47	18.48	0.063	0.17	0.18	0.14	0.15
134	Dichlorofluoromethane	Training	19.90	20.63	20.67	0.053	0.73	0.77	0.59	0.63
135	Dichlorofluoropropane	Training	23.00	21.68	21.66	0.015	-1.32	-1.34	-1.04	-1.06
136	Dichlorohexafluorocyclobutane	Training	19.30	19.51	19.52	0.028	0.21	0.22	0.17	0.18
137	Dichloromethane	Training	23.10	22.38	22.34	0.046	-0.72	-0.76	-0.58	-0.61
138	Dichlorotetrafluoro propane	Training	19.10	19.06	19.06	0.010	-0.04	-0.04	-0.03	-0.03
139	Dichlorotrifluoropropane	Training	19.70	20.04	20.04	0.009	0.34	0.34	0.27	0.27
140	Diethyl ether	Training	19.09	18.01	17.94	0.063	-1.08	-1.15	-0.87	-0.93
141	Difluoromethane	Training	6.90	7.71	7.81	0.118	0.81	0.91	0.67	0.76
142	Difluoromethyl2,2,2-trifluoroethyl ether	Training	12.60	13.54	13.58	0.037	0.94	0.98	0.75	0.78
143	Difluoropropane	Training	12.40	13.32	13.34	0.021	0.92	0.94	0.73	0.75
144	Ethane	Training	10.86	12.57	12.62	0.026	1.71	1.76	1.36	1.40
145	Ethylamine	Training	21.60	20.06	19.82	0.131	-1.54	-1.78	-1.30	-1.50
146	Fluoroethane	Training	10.70	11.35	11.36	0.021	0.65	0.66	0.52	0.53
147	Fluoroethylene	Training	9.90	12.68	12.74	0.019	2.78	2.84	2.21	2.25
148	Fluoromethane	Training	6.70	9.55	9.95	0.121	2.85	3.25	2.39	2.72
149	Fluoropropane	Training	13.80	13.84	13.84	0.028	0.04	0.04	0.03	0.03
150	Hexachloroethane	Training	42.80	43.21	43.25	0.084	0.41	0.45	0.34	0.37
151	Hexafluoroethane	Training	8.00	6.39	6.34	0.032	-1.61	-1.66	-1.28	-1.33
152	Hexafluoropropylene	Training	9.60	9.65	9.66	0.022	0.05	0.06	0.04	0.04
153	Isobutane	Training	15.60	17.54	17.68	0.067	1.94	2.08	1.58	1.69
154	Methyl1,1,2,2-tetrafluoroethyl ether	Training	12.70	12.68	12.68	0.027	-0.02	-0.02	-0.01	-0.01

155	Methyl formate	Training	21.30	15.32	14.77	0.084	-5.98	-6.53	-4.91	-5.36
156	Methylpentafluoroethyl ether	Training	12.20	12.40	12.40	0.030	0.20	0.20	0.16	0.16
157	Methylamine	Training	18.40	17.68	17.60	0.100	-0.72	-0.80	-0.59	-0.66
158	Octafluorocyclobutane	Training	11.00	10.16	10.13	0.042	-0.84	-0.87	-0.67	-0.70
159	Octafluoropropane	Training	9.30	8.19	8.16	0.028	-1.11	-1.14	-0.88	-0.91
160	Pentachloroethane	Training	38.60	37.48	37.42	0.049	-1.12	-1.18	-0.90	-0.95
161	Pentafluorodimethyl ether	Training	10.30	10.08	10.05	0.108	-0.22	-0.25	-0.19	-0.21
162	Pentafluoroethane	Training	8.30	7.41	7.39	0.025	-0.89	-0.91	-0.71	-0.72
163	Pentafluoropropane	Training	10.30	10.84	10.85	0.013	0.54	0.55	0.43	0.43
164	Pentane	Training	18.70	19.73	19.85	0.102	1.03	1.15	0.85	0.95
165	Propane	Training	14.20	14.71	14.74	0.039	0.51	0.54	0.41	0.43
166	Propylene	Training	13.62	15.46	15.52	0.032	1.84	1.90	1.47	1.52
167	Tetrachlorodifluoro propane	Training	29.80	28.80	28.77	0.025	-1.00	-1.03	-0.80	-0.82
168	Tetrachlorotrifluoro propane	Training	27.40	27.61	27.62	0.025	0.21	0.22	0.17	0.17
169	Tetrafluoroethylene	Training	8.10	8.63	8.64	0.029	0.53	0.54	0.42	0.43
170	Tetrafluoromethane	Training	6.00	3.75	3.37	0.143	-2.25	-2.63	-1.91	-2.23
171	Trichlorodifluoro propane	Training	25.10	24.61	24.60	0.014	-0.49	-0.50	-0.39	-0.39
172	Trichloroethylene	Training	31.00	33.60	33.70	0.037	2.60	2.70	2.08	2.16
173	Trichlorofluoropropane	Training	26.80	25.89	25.88	0.018	-0.91	-0.92	-0.72	-0.73
174	Trichloromethane	Training	28.90	29.39	29.43	0.077	0.49	0.53	0.40	0.44
175	Trichlorotetrafluoro propane	Training	22.30	23.05	23.06	0.016	0.75	0.76	0.59	0.60
176	Trichlorotrifluoro propane	Training	24.20	24.20	24.20	0.015	0.00	0.00	0.00	0.00
177	Trifluoromethyl1,1,2,2-tetrafluoroethyl ether	Training	11.25	11.10	11.09	0.044	-0.15	-0.16	-0.12	-0.13
178	Trifluoromethyl1,2,2,2-tetrafluoroethyl ether	Training	11.20	11.53	11.55	0.049	0.33	0.35	0.27	0.28
179	Trifluoropropane	Training	11.80	12.30	12.31	0.016	0.50	0.51	0.40	0.40
180	1,1,1,2,2,3,3-Heptafluorobutane	Test	11.22	-	11.50	0.015	-	0.28	-	-
181	1,1,2,3Pentafluoro propane	Test	10.70	-	10.80	0.013	-	0.10	-	-
182	1,1,1,3,3,3Hexafluoro propane	Test	10.30	-	10.31	0.015	-	0.01	-	-
183	1,1,1,3-Tetrachloro-2,2,3-trifluoropropane	Test	27.40	-	27.66	0.026	-	0.26	-	-
184	1,1,1-Trichloro-2,2-	Test	24.60	-	24.80	0.013	-	0.20	-	-

	difluoroethane									
185	1,1,2,2,3,3,3- Heptafluoropropane	Test	9.60	-	9.20	0.020	-	-0.40	-	-
186	1,1,2,2-Tetrafluoroethane	Test	8.70	-	7.97	0.021	-	-0.73	-	-
187	1,1,2-Trichloro-2,2- difluoroethane	Test	24.60	-	24.65	0.013	-	0.05	-	-
188	1,1,3,3-Tetrachloro-1,2,2- trifluoropropane	Test	27.40	-	27.65	0.026	-	0.25	-	-
189	1,1-Dichloro-1,2,2,3,3- pentafluoropropane	Test	17.60	-	18.46	0.013	-	0.86	-	-
190	1,1-Dichloro-1,2,3,3,3- pentafluoropropane	Test	17.60	-	18.47	0.013	-	0.87	-	-
191	1,1-Dichloro-2,2,3,3- tetrafluoropropane	Test	18.50	-	18.95	0.009	-	0.45	-	-
192	1,1-Dichloro-2-fluoroethane	Test	22.10	-	20.25	0.011	-	-1.85	-	-
193	1,2,2,2-Tetrafluoroethyl difluoromethyl ether	Test	11.70	-	11.71	0.038	-	0.01	-	-
194	1,2-Dichloro-1,1,2,3,3,3- hexafluoropropane	Test	16.80	-	17.39	0.019	-	0.59	-	-
195	1,2-Dichloro-1,1- difluoroethane	Test	20.50	-	19.20	0.010	-	-1.30	-	-
196	1,2-Dichloroethane	Test	25.00	-	21.16	0.016	-	-3.84	-	-
197	1,3-Dichloro-1,1,2,2,3- pentafluoropropane	Test	17.60	-	18.50	0.013	-	0.90	-	-
198	1,3-Dichloro-2,2- difluoropropane	Test	22.16	-	20.65	0.011	-	-1.51	-	-
199	1-Chloro-1,1,2,2- tetrafluoropropane	Test	14.70	-	15.11	0.009	-	0.41	-	-
200	1-Chloro-1,1,3,3,3- pentafluoropropane	Test	14.40	-	14.54	0.010	-	0.14	-	-
201	1-Chloro-1,2,2,3- tetrafluoropropane	Test	15.10	-	15.00	0.009	-	-0.10	-	-
202	1-Chloro-1fluoropropane	Test	18.20	-	17.47	0.019	-	-0.73	-	-
203	2,2-Dichloro-1,1,1,3,3- pentafluoropropane	Test	17.60	-	18.43	0.012	-	0.83	-	-

204	2,3-Dichloro-1,1,1,2,3-pentafluoropropane	Test	17.60	-	18.39	0.012	-	0.79	-	-
205	2-Chloro-1,1,1,3,3,3-hexafluoropropane	Test	13.40	-	13.42	0.015	-	0.02	-	-
206	3,3-Dichloro-1,1,1,2,2-pentafluoropropane	Test	17.60	-	18.43	0.012	-	0.83	-	-
207	3-Chloro-1,1,2,2-tetrafluoropropane	Test	15.10	-	15.08	0.009	-	-0.02	-	-
208	Bromochlorodifluoromethane	Test	23.10	-	23.07	0.077	-	-0.03	-	-
209	Chlorodifluoroethane	Test	14.70	-	14.79	0.011	-	0.09	-	-
210	Chlorofluoromethane	Test	14.00	-	13.57	0.035	-	-0.43	-	-
211	Chloropentafluoroethane	Test	12.30	-	11.49	0.020	-	-0.81	-	-
212	Chlorotrifluoropropane	Test	15.50	-	16.01	0.010	-	0.51	-	-
213	Dichlorodifluoropropane	Test	20.90	-	20.54	0.010	-	-0.36	-	-
214	Dichloropentafluoro propane	Test	17.60	-	18.43	0.012	-	0.83	-	-
215	Difluoromethyl,1,2-trifluoroethyl ether	Test	12.61	-	12.55	0.031	-	-0.06	-	-
216	Ethylene	Test	9.68	-	14.17	0.024	-	4.49	-	-
217	Hexachlorodifluoro propane	Test	36.30	-	36.88	0.070	-	0.58	-	-
218	Isopentane	Test	17.60	-	20.03	0.105	-	2.43	-	-
219	Methyltrifluoromethyl ether	Test	11.40	-	11.98	0.086	-	0.58	-	-
220	Pentachlorofluoroethane	Test	35.90	-	36.49	0.048	-	0.59	-	-
221	Perfluoropropylmethyl ether	Test	12.70	-	13.55	0.029	-	0.85	-	-
222	Tetradecafluorohexane	Test	11.30	-	9.60	0.034	-	-1.70	-	-
223	Trichlorofluoromethane	Test	26.30	-	27.84	0.082	-	1.54	-	-
224	Trifluoromethane	Test	6.20	-	5.90	0.125	-	-0.30	-	-

III-2 Développement du modèle

L'équation de régression du modèle calculé est la suivante :

$$Ts = -25.8 + 52.5 Mp + 4.57 H1m + 53.6 RDF0\ 10m - 5.87 MATAS1v \quad (14)$$

Tableau IV: Signification des descripteurs

Descripteurs	Signification
Mp	Point de fusion
H1m	Autocorrélation du décalage 1/ pondérée par masse atomique
RDF0 10m	Fonction de distribution radiale- 10/ pondérée par masse atomique
MATAS1v	Autocorrélation de Moran-décalage 1/pondérée par masse atomique

Les paramètres statistiques sont présentés dans le tableau suivant :

Tableau V: Les paramètres statistiques pour l'ensemble de calibration (179 composés) et celui de validation (45 composés)

Size	R2	Q2 _{loo}	Q2 _{boot}	Q2 _{ext}	R2 _{adj}	LOF	AIC	EQMC	EQMP	Kx	Kxy
	97.03	96.76	96.67	96.77	96.96	1.727	1.714	1.255	1.311	37.94	49.20

Size	F	SE	DF	DK	DQ	PRESS	EQMP _{ext}
	1420.16	1.273	174	0.113	-0.001	96.78	1.867

Les paramètres statistiques reproduits dans le tableau ci-dessus montrent une bonne explication de la variabilité Ts. Le coefficient de détermination, et une robustesse du modèle due à la valeur de Q^2_{LOO} , alors que les valeurs de l'erreur quadratique moyenne de prédiction et de calcul sont petits (par rapport aux valeurs de Ts) et proches; en outre ce modèle est significatif avec une valeur du paramètre de Fisher ($F = 1420.16$), Les statistiques montrent aussi que le modèle à une très grande capacité prédictive externe, ceci traduit par la grande valeur de Q^2_{ext} et la faible valeur de EQMP_{ext}.

III-2-1 Qualité de l'ajustement

La qualité de l'ajustement a été vérifiée par le graphe des valeurs calculées de l'ensemble de calibration et les valeurs prédites de l'ensemble de validation en fonction de celles expérimentales du la tension superficielle

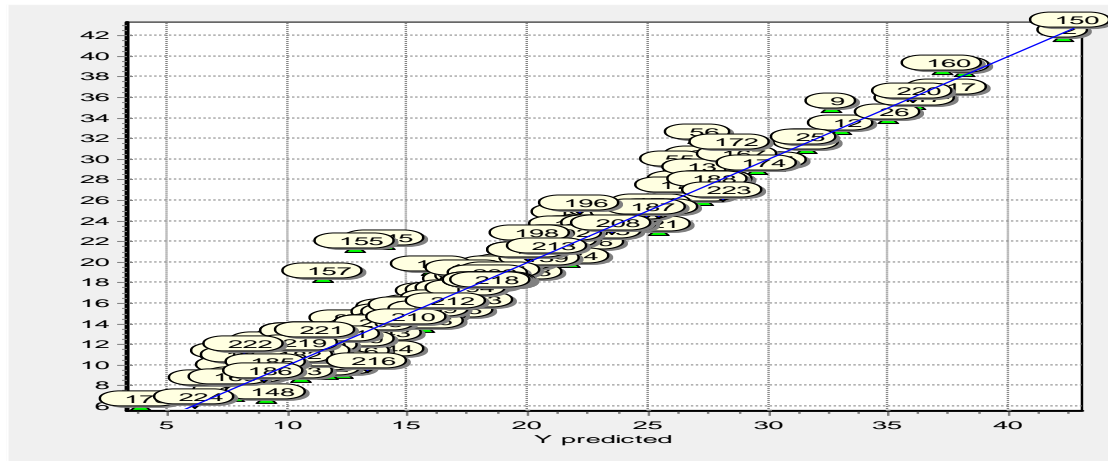


Figure 05: Qualité de l'ajustement

D'après le graphe on remarque une bonne dispersion autour de la première bissectrice des deux ensembles de calibration et de validation.

Ceci est confirmé par la valeur de Q^2_{Loo} , élevée (=96.77 %).

III-2-2 Diagramme de Williams

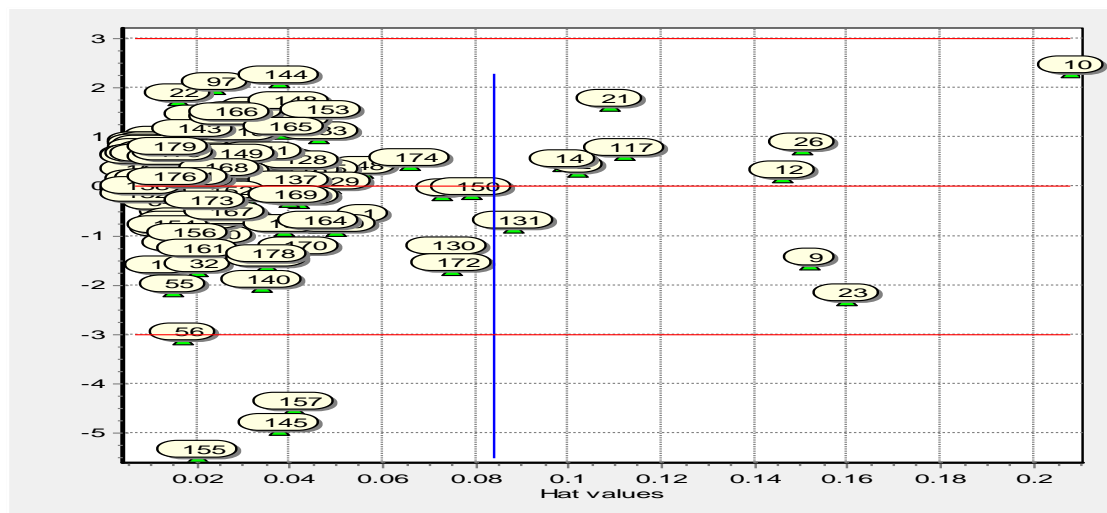


Figure04:Diagramme de Williams

D'après le graphe on remarque qu'ils existent trois composés (Ethylamine, Methyl formate, Methylamine) qui sont inférieure aux limites ± 3 , ce sont des points influents négative sur notre modèle et (10) autre composés ce sont des points aberrants car sa valeur de levier (h_{ii}) et très éloigner par rapport h^*_c c'est-à-dire 166 sont les points ont une valeur de levier inférieur à la valeur critique $h^*=0.083$ $h^*=(3(k+1))/n$.

III-2-3 Test de randomisation

Dans le but de montrer que le modèle obtenu n'est pas dû au hasard nous avons appliqué le test de randomisation de Y , dans cette figure on a présenté le graphe des coefficients statistiques Q^2 et R^2 .

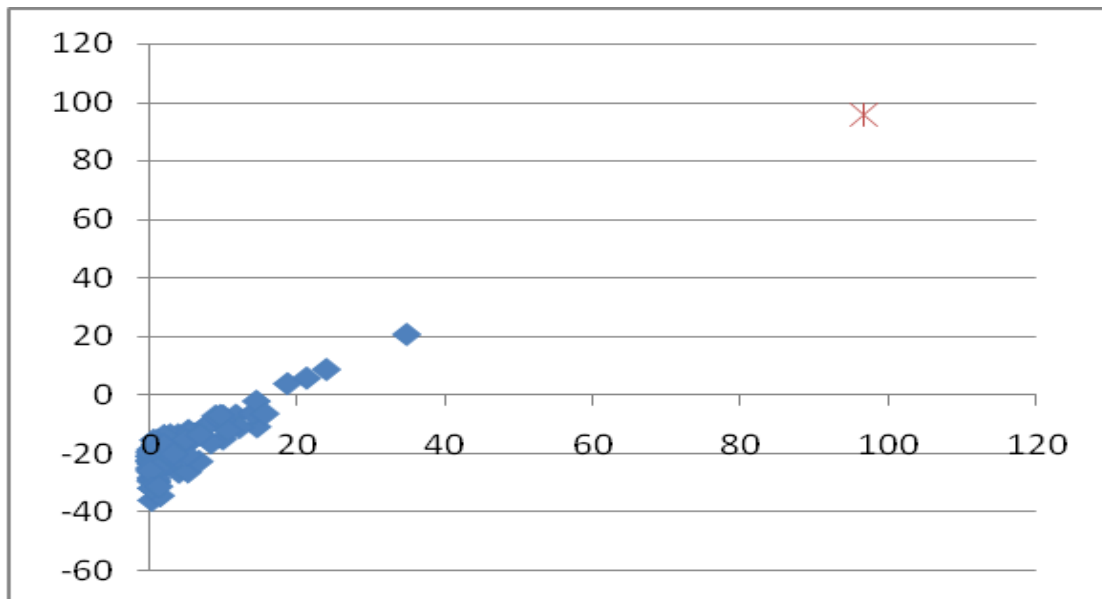


Figure 6: Test de randomisation

Il est clair que les statistiques obtenues pour les vecteurs TS modifiés sont plus petites que celles du modèle réel, les losanges vides regroupés dans

la région des valeurs négatives de Q_2 ont des valeurs petites de R_2 , seul le cercle noirci a des valeurs élevés et proches pour ces deux statistiques, il représente notre modèle calculé, donc il n'est pas dû au hasard.

CONCLUSION GENERALE

Conclusion

Nous avons utilisé la méthodologie QSPR pour relier la tension superficielle des dérivées de propane avec des descripteurs moléculaires théoriques caractéristique de la molécule entière ou de ses fragments, calculés à l'aide de logiciels spécialisés du commerce.

Nous avons recherché des corrélations linéaires entre la variable dépendante et les variables explicatives sélectionnées par algorithme génétique, en utilisant la régression linéaire multiple.

Le modèle obtenu conduit aux bons résultats : qualité de l'ajustement, robustesse interne et externe, capacité prédictive ... ce qui prouve que la relation entre tension superficielle et la variables explicatives choisis elle est fondamentalement linéaire.

D'après l'étude qui nous allons faire et les résultat des autre traveaux précédent , on remarque que toujours le point de fusion sélectionné pour l'explication de la tension supperficielle (T_s) ce qui confirme l'influence de cette paramètre sur T_s .

REFERENCES BIBLIOGRAPHIQUES

- [1] DEFAY R. et prigogine I ; Tension superficielle et adsorption ; Editions DESOER à LIEGE Belgique. 1952.
- [2] S Chtita, Modélisation de molécules organique hétérocyclique biologiquement actives par des méthodes QSAR/QSPR. Recherche de nouveaux médicaments, (2017)
- [3] CAL Horstmann, FEC Scheffer (1932). Recueil des Travaux chimiques des pays-bas. Vol 51, N 2 Pages, 98-217.
- [4] F Cyrot-Lackmann (1968). Sur le calcul de la cohésion et de la tension superficielle des métaux de transition par une méthode de liaisons fortes, *Journal of Physics and Chemistry of Solids*. Vol 29 issue 7 Pages 1235-1243.
- [5] Pierre-Marie GASSIN 2014 *Mesure de la tension superficielle par la technique de la goutte pendante*. Union des professeurs de physique et de chimie.
- [6] Ramachandran K.I , Deepa, G , Namboori, K. *Computational Chemistry and Molecular Modeling. Principles and Applications* 2008. DOI 10. 1007/978-3-540-77304-7.
- [7] Dodge Y., Rousson. (2004)- *Annalyse de regression appliqué* Dunold, Paris.
- [8] Montgomery, D.C , Peck, E.A. (1992)-*Introduction to Linear Regression Analysis*. 5nd Edition. John Wiley & sons. Inc.
- [9] Cristianini N., Shawe-Taylor J. (2000)- *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press.
- [10] Kubinyi H, (1994)- *Quantitative Structure-activity Relationships...*, Vol 13 , pp 285.

- [11] Draper N.R, and Smith H., (1998)- Applied Regression Analysis, Third Edition, Wileyseries in Probability and Statistics, New York.
- [12] Eriksson L., Jaworska J, and Worth A.P., (2003)- Perspective M.T.D., Vol 111(10),pp****1361-1375.
- [13] Erikson L., Jaworska J., Worth A. P., Cronin M. T. D., Mc Dowell R. M, and Gramatica P., (2003) - Methods for Reliability and uncertainty Assessment and for Applicability Evaluations of Classification-and Regression - Based QSARs. *J Environmental Health Perspectives*. Vol 111 (10). pp 1361-1375.
- [14] Tropsha A, Gramatica P, and Grombar V.K. (2003)- The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR & J Combinatorial Science.*, Vol 22. pp 69-76.
- [15] Sharma B.K, Singh P, Pilania P, Sarbhai K, Yenamandra S, and Prabhakar. CP-(2011) - MLR/PLS directed QSAR study on apical sodium-codependent bile acid transporter inhibition activity of benzothiepinines *J Mol Divers* Vol 15. pp 135–147.
- [16] A brpgar, khageh, Hamid. (2012)- M journal international journal of rener Vol 35. pp 150 -159
- [17] HYPERCHEM Software (2002), Release 6.03 for Windows, Molecular Modeling System.
- [18] R. Todeschini , V.Consonni , A. Mauri , M. Pavan, (2005). Dragon Software for the calculation of molecular descriptors , Version 5.3 for Windows , Taletesrl, Milano . [Italy](#)
- [19] R. Leardi, R. Boggia , M. Terrile (1992)., Genetic algorithms as a strategy for featureselection, *J. Chemom*, **6**, 267

