

## Département Génie Chimique

### Projet fin d'étude

Présenté pour l'obtention du titre d'ingénieur de l'Ecole Nationale Polytechnique  
en Génie Chimique

#### ● ● THEME ● ●

**CARACTERISATION DES FRACTIONS PETROLIERES PAR LE  
DEVELOPPEMENT D'UN NOUVEAU MODELE NUMERIQUE BASE SUR  
LES ALGORITHME GENETIQUE**

Proposé et dirigé par :

■ Mme Dr. F.SOUAHI

Etudié par :

■ BELGHIT Chafik

**PROMOTION JUIN 2006**

## Dédicaces

*Aux les seuls deux personnes qui ont partagé avec moi les moments d'inquiétude avant ceux de joie, pendant ces longues années de ma vie. Leurs soutiens et leurs avis m'ont toujours aidé à dépasser les moments difficiles. Il n'y pas de mots suffisamment forts pour leurs exprimer ma gratitude. Je fait remercie énormément pour la compréhension et la patience dont ils ont preuve durant ces années. Mes parents, qui ont toujours cru en moi et qui, malgré la distance, m'ont accompagné moralement tout au long de ce parcours. L'éducation, soins et conseils qu'ils m'ont prodigués, m'ont toujours guidé pour faire les bons choix dans ma vie. Ils m'ont appris comment réaliser les rêves difficiles. Sans leurs soutiens, encouragements et prières, rien n'aurait été possible.*

*A mon frère et toutes mes soeurs qui, malgré la distance étaient toujours là pour me conseiller, me soutenir, m'encourager et prier pour moi. Je les remercie de tout mon coeur.*

*A mes oncles et tantes et tout particulièrement à Belghit Abdérahim.*

## Remerciements

Tâche ô combien délicate que celle qui consiste à rédiger ces premières pages ! Ici plus que nulle part ailleurs dans un mémoire, le choix des mots et l'ordonnancement des idées revêtent une importance démesurée, et les règles et phrases toutes faites de nos prédécesseurs, loin de simplifier les choses, ne font qu'imposer à l'auteur un périlleux exercice de style où il doit faire preuve d'originalité sans toutefois avoir le droit à l'erreur.

Qu'il me soit d'abord permis de remercier et d'exprimer ma gratitude envers **Allah**, qui m'a donné la patience et le courage pour que je puisse continuer ce travail.

Je tiens également à remercier Madame **Dr. Fatiha. Souahi**, Maître de Conférences à l'ENP (Ecole Nationale polytechnique) d'Alger, pour avoir dirigé ce travail et avec qui j'ai beaucoup appris dans les domaines théorique et pratique. Je la remercie de tout mon cœur pour le temps qu'elle m'a consacré ainsi que pour les nombreuses discussions scientifiques.

Je remercie vivement Madame **Faroudja Mohellebi** et Monsieur **Rachid Illoul**, pour avoir accepté de participer à la critique de ce travail en siégeant dans ce jury.

Je ne veux pas oublier, dans mes remerciements, les membres de département de Génie Chimique. Mes remerciements vont notamment et plus particulièrement à :

Madame **Derriche** et Madame **Cherchari** et Monsieur **Pr-Chitour** et Monsieur **Slatnia** et Monsieur **Benyoussef** et Monsieur **Ahmed Zaid** pour leurs enseignements et leurs gentillesse tout au long des années passées au sein du département du Génie Chimique.

Je ne peux oublier, bien sûr, tous mes amis de la résidence universitaire **Bouraoui Amar**, avec qui j'ai passé les meilleurs moments durant ces dernières années et dont les conseils m'ont beaucoup aidé pour continuer mon travail dans les meilleures conditions.

La bonne humeur et l'ambiance au sein du groupe m'ont permis de mener mes travaux de recherche de façon agréable. Les discussions, les remarques et les commentaires de mes collègues ont été source d'idées et ont contribué au développement et à l'amélioration de cette recherche.

Je voudrais vivement remercier mes amis (es) et mes cousins et cousines. **Abdallah, Abdekader, Abdnasser, Ali, Amina, Ahlem, billal, badi, larbi, Fateh, Fathi, Mounir, Nawfel, Nadhir, Nouar, NourEdine, NourEzmen, Khaoula, Saber, Sabri, Salim, Soufiane, Ramzi, Zakaria, Zouhir**, avec qui j'ai toujours partagé les moments d'inquiétude et ceux de joie.

Enfin, je tiens à remercier vivement une personne que j'apprécie beaucoup, **Hamidane H'mida**. Les paroles de cet ami m'ont toujours encouragé et m'ont toujours donné l'envie d'être meilleur.

# Sommaire

Introduction générale .....	1
Chapitre I : Généralités sur le pétrole	
I-1 Introduction .....	3
I-2 Etat physique .....	4
I-3 La formation du pétrole .....	4
I-4 La problématique des réserves .....	5
I-4 Production .....	8
I-5 Consommation .....	10
I-6 L'évolution des prix du pétrole .....	11
I-7 Secteurs d'utilisation .....	11
Chapitre II : Etude des propriétés physiques et thermodynamiques	
II-1 Introduction .....	14
II-2 Les propriétés critiques .....	15
II-2-1 Mesure expérimentale des paramètres critiques .....	15
II-2-2 Estimations des propriétés critiques des fractions pétrolières .....	16
II-2-2-1 Température critique des fractions pétrolières .....	16
II-2-2-2 Pression critique des fractions pétrolières .....	17
II-2-2-3 Volume critique des fractions pétrolières .....	18
II-3 Facteur acentrique .....	19
Chapitre III : Les Algorithmes Génétiques	
III-1 Introduction .....	21
III-1-1 Définition d'un individu .....	21
III-1-2 Codage des variables .....	22
III-1-3 Définition d'une population .....	22
III-2 Déroulement de l'algorithme .....	23
III-2-1 Les opérations .....	24
a- La Reproduction / Sélection .....	24
a- 1 L'élitisme .....	25
a- 2 Le tournoi .....	25
a-3 La roue biaisée .....	25
b. Les croisements .....	26
b-1 Le croisement uniforme .....	26
b-2 Le croisement à une ou deux coupures .....	26

## Sommaire

c- La mutation .....	27
III-3 Prise en compte des contraintes par pénalisation .....	28
III-4 Convergence et inconvénients des algorithmes génétiques .....	29
4-1. Inconvénients .....	29
4-2. Critère de convergence .....	30
Chapitre IV : Application et programmation des AG	
IV-1 Enoncé du problème .....	31
IV-1-1 Choix du modèle de chaque propriété .....	31
IV-1-2 Détermination de l'espace de recherche des coefficients de chaque modèle .....	31
IV-2 L'organigramme de la programmation des A.G .....	32
IV-2-1 Codage réel des paramètres .....	32
IV-2-2 Création de la population initiale .....	33
IV-2-3 Calcul du fitness .....	34
IV-2-4 Opérateur de sélection .....	34
IV-2-5 Opérateur de croisement .....	35
IV-2-6 Opérateur de Mutation .....	36
IV-2-7 Test d'arrêt .....	38
IV-3 Exemple mettant en œuvre toutes les opérations .....	38
<b>a-</b> Tirage de la population initiale .....	38
<b>b-</b> population sélectionnée par la reproduction .....	39
<b>c-</b> population après l'opération de croisement .....	39
<b>d-</b> population après l'opération de mutation .....	40
Chapitre V : Résultats et commentaires	
V.1 Introduction .....	41
V.2 Résultats obtenus .....	41
V.2-1 Température critique .....	41
V.2-1-2 Etude graphique .....	43
V.2-2-1 Pression critique .....	45
V.2-2-2 Etude graphique .....	47
V.2-3-1 Volume critique .....	49
V.2-3-2 Etude graphique .....	51
V.2-4 Commentaires .....	53
V.2- 5-1 Facteur acentrique .....	53
V.2-5-2 Etude graphique .....	55

## Sommaire

V.2-6 Commentaire .....	57
V-3 Conclusion .....	57
<b>Conclusion générale</b> .....	58
Rréférence bibliographique.....	59
<b>Annexe</b>	

## Introduction générale

L'une des rares certitudes sur lesquelles toutes les analyses des experts s'accordent, est que les besoins énergétiques mondiaux continueront de s'accroître à un rythme soutenu les prochaines années et que le pétrole restera la source d'énergie dominante. Aussi, sa mise en œuvre optimale dans l'industrie du raffinage, le dimensionnement des équipements dans les unités de fabrication et la simulation des procédés et leur optimisation, reposent sur une connaissance parfaite des caractéristiques physico-chimiques et thermodynamiques de ce pétrole. L'expérience fournit la majorité de ces caractéristiques mais les limitations qu'elle présente telles que la complexité de la mise en œuvre, le coût et la perte de temps font que des méthodes de calcul en été élaborées pour estimer de façon plus ou moins précise ces paramètres.

Pour les mélanges complexes de type fraction pétrolière, il n'est pas toujours possible d'analyser avec précision leurs caractéristiques pour en déduire la nature et la composition de chaque constituant. Aussi, pour pouvoir représenter de tels mélanges, la notion de pseudo-constituant est introduite. Ces derniers jouent, au sein d'une coupe pétrolière, un rôle équivalent à celui d'un corps pur dans un mélange défini. Il est alors nécessaire, pour tout calcul thermodynamique, de connaître les propriétés intrinsèques de ces pseudo-constituants. Ces calculs sont possibles grâce à l'utilisation de corrélations spécifiques ne nécessitant pas la connaissance de la structure chimique du pseudo-constituant. La génération des propriétés de pseudo-constituants peut être effectuée en fournissant pour chaque pseudo-constituant les données de base nécessaires pour les corrélations utilisées. Il convient donc de fournir au moins deux propriétés caractéristiques parmi les suivantes :

- température d'ébullition sous 1 atm,
- densité spécifique (SPGR) ou degré API,
- facteur de caractérisation de WATSON,
- masse molaire.

Pour aller plus loin dans ces travaux, l'étude qui nous a été confiée dans le cadre de cette thèse a pour objectif principal la mise au point de nouveaux modèles numériques pour la prédiction de quatre propriétés telle que les trois propriétés critiques et le facteur acentrique des fractions pétrolières.

Nous avons utilisé une nouvelle méthode de calcul des propriétés des fractions pétrolières par procédure de caractérisation basée sur l'utilisation de l'intelligence artificielle : les Algorithmes Génétiques (AG).

## Introduction générale

Le présent rapport de thèse comporte cinq chapitres :

Le chapitre I expose d'une manière succincte les généralités sur le pétrole brut. Une description de cette matière première d'un point de vue composition chimique, production, consommation, variation du prix, secteur d'utilisation et aussi donnée.

Le chapitre II est scindé en deux parties :

Dans la première partie nous présentons, quelques concepts de base de la thermodynamique aux propriétés que nous allons étudier.

Dans la deuxième partie, nous abordons le problème de la détermination des propriétés physiques et thermodynamiques des fractions pétrolières. Pour chaque propriétés étudiées  $(T_{pc}, P_{pc}, V_{pc}, \omega)$ , une définition ainsi qu'un fondement thermodynamique sont données.

Les méthodes exprimables, quand elles existent, sont également présentées, suivies des méthodes de calcul disponibles dans la littérature qui permettent de prédire ces paramètres rapidement dans le cas des fractions pétrolières.

Dans le chapitre III nous avons présenté les Algorithmes Génétiques

Le chapitre IV : nous avons développé la méthode de calcul et de programmation des AG pour caractériser les fractions pétrolières.

Le chapitre V regroupe les résultats obtenus dans le cadre de ce travail avec une étude comparative avec d'autres corrélations trouvées dans la littérature.



# Chapitre I :

# Généralités sur le pétrole

## I-1 Introduction :

Le pétrole est un produit constitué d'une multitude de molécules composées d'atomes de carbone et d'hydrogène uniquement, on parle d'hydrocarbures.

Ces différentes molécules sont caractérisées par le nombre et la structure des atomes de carbone. La chimie minérale distingue quatre grands types d'hydrocarbures en fonction de la nature des liaisons atomiques entre atomes de carbone:

- les alcanes : ce sont les molécules dont la chaîne d'atomes de carbone est composée de liaisons simples, c'est une chaîne aliphatique saturée. Le nombre de carbone varie le plus souvent entre 1 et 20. Autrefois, les alcanes étaient appelés paraffines. Exemples : le méthane, le butane, l'hexane. Formule brute :  $C_nH_{2n+2}$ .

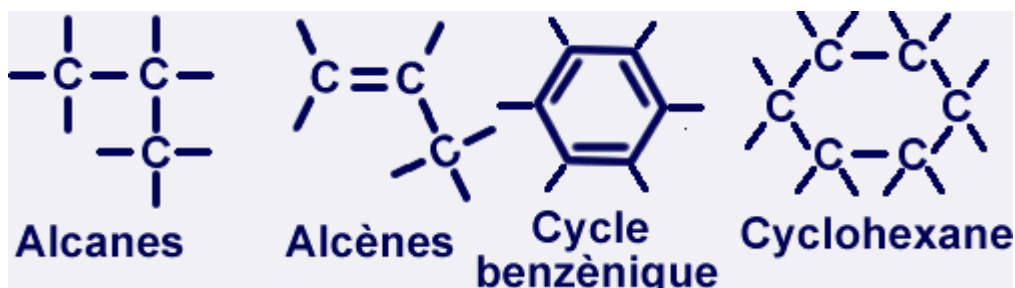
Un groupe d'alcanes est particulier, ce sont les cycloalcanes dont la chaîne d'atomes de carbone contient au moins un cycle hexane ou pentane. Exemples : le cyclohexane, le méthylcyclopentane. Formule brute :  $C_nH_{2n}$

- les alcènes ou carbures éthyléniques : ce sont les molécules dont la chaîne d'atomes de carbone contient une liaison double. Le nombre de carbone varie le plus souvent entre 2 et 20. Autrefois, les alcènes étaient appelés oléfines. Exemples : l'éthylène, le butène, l'isobutène. Formule brute :  $C_nH_{2n}$

- les diènes ou alcynes : ce sont les molécules qui contiennent plus d'une double liaison dans la chaîne aliphatique. Exemples : l'acétylène, les butadiènes.

Formule brute :  $C_nH_{2n-2}$

- les chaînes aromatiques : ce sont les molécules dont la chaîne d'atomes de carbone contient au moins un cycle benzénique. Le nombre de carbone est au minimum de 6. Exemples : le benzène, le naphthalène. Formule brute :  $C_6H_5-Y$  (où Y représente une molécule attachée au cycle benzénique). [1], [2]



## **I-2 Etat physique :**

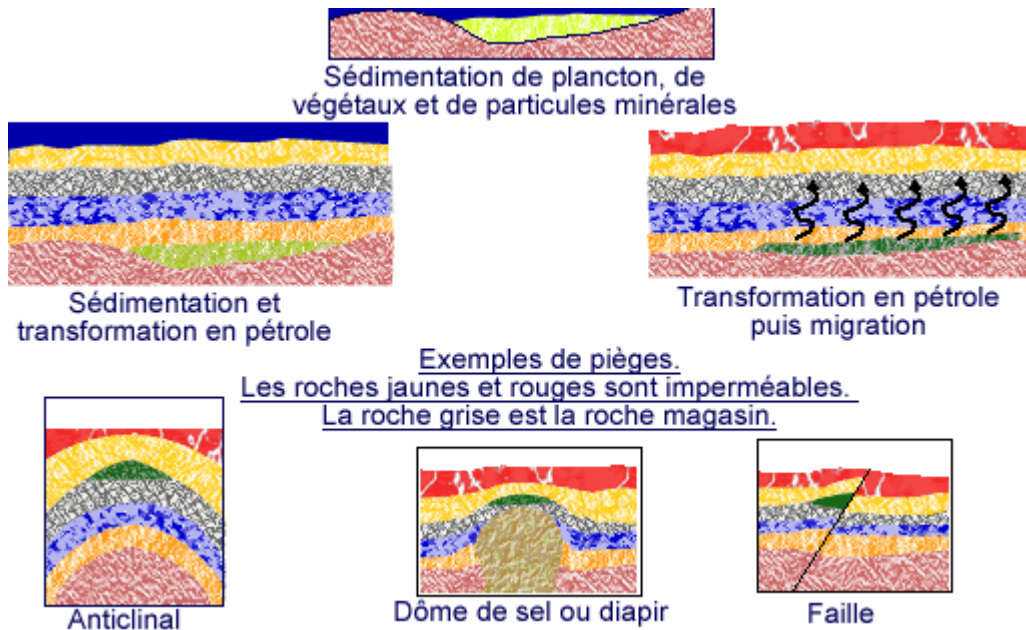
L'état physique des composés du pétrole dans les conditions normales varie suivant la longueur de la chaîne aliphatique. En effet, plus la chaîne est longue, plus la masse volumique du composé est grande et plus le composé est visqueux. A titre d'exemple, les composés du méthane au butane (C1 à C4) sont gazeux à température ambiante ; à partir de cinq atomes de carbone les composés sont liquides et au delà d'une quinzaine d'atomes de carbone, ils sont solides. Comme pour toute molécule, l'état physique des composés varie en fonction de la température et de la pression. Ces propriétés sont utilisées en vue de séparer les chaînes de différentes longueurs : la distillation est un exemple de séparation des molécules par la chaleur.[3]

## **I-3 La formation du pétrole :**

Le pétrole (littéralement huile de pierre) est issue de la décomposition de matières organiques végétales et animales. Ces particules se déposent en même temps qu'une sorte de boue. Le milieu d'accumulation est confiné : lac, lagunes, deltas ou lagons. Ce milieu étant peu oxygéné, des réactions réductrices transforment la matière organique en kérogène. Au fur et à mesure du recouvrement de cet ensemble boue - kérogène, la transformation en hydrocarbures commence. Elle débute vers 60°C, ce qui correspond à un enfouissement d'environ 1500 à 2000 mètres. La vitesse de transformation augmente jusqu'à 100°C (3000m) puis, au-delà de cette limite, diminue lorsque la température augmente. Si l'enfouissement est supérieur à 4000m (soit plus de 150°C), il ne se forme plus que du méthane. En tout état de cause, ces transformations sont très lentes et nécessitent plusieurs millions d'années.

La matière organique en décomposition est en quelque sorte coincée dans la roche qui s'est formée au moment de son enfouissement. On parle de roche mère. Le pétrole en formation est moins dense que la roche qui l'entoure, il peut donc migrer vers la surface à condition qu'il ne soit pas arrêté dans sa progression par des roches imperméables par exemple. Si tel est le cas, le pétrole est alors stoppé dans sa remontée et se concentre pour former des poches. Ces pièges sont à l'origine des réservoirs actuels de pétroles.[2], [3]

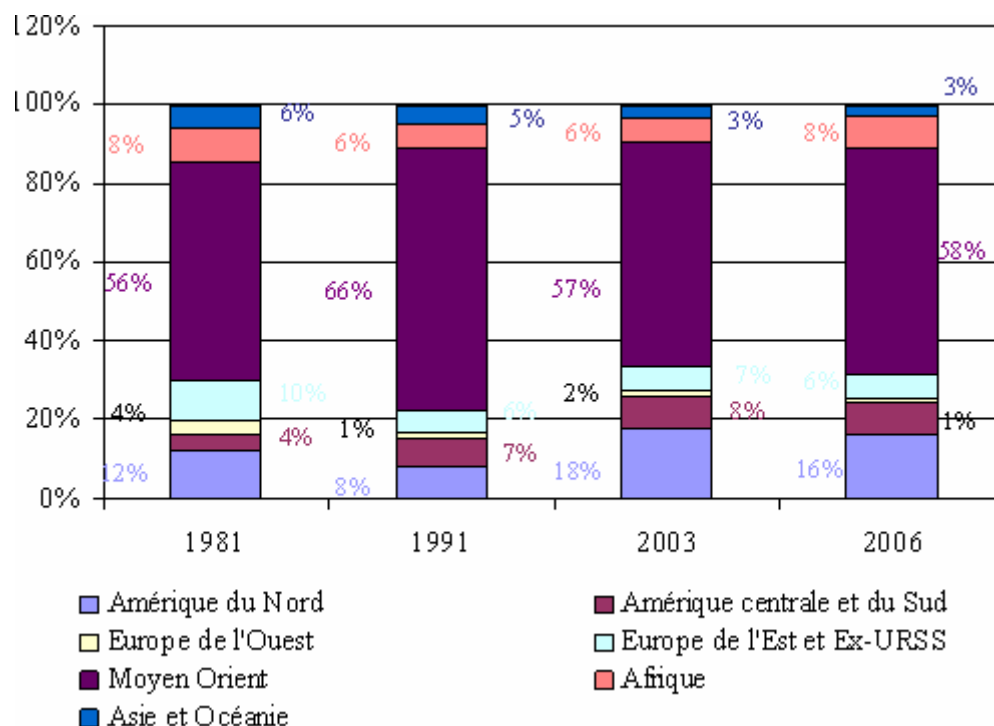
### Etapes de la formation du pétrole et quelques exemples de pièges



## I-4 La problématique des réserves

Evaluer les réserves de pétrole est une question importante afin de gérer au mieux cette ressource non renouvelable. En effet, une fois l'ensemble des ressources pétrolifères épuisées, cette source d'énergie sera définitivement tarie. Cependant les réserves évoluent sans cesse, et contrairement à ce qu'on pourrait penser les réserves identifiées (ou prouvées) n'ont jamais été aussi importantes qu'à l'heure actuelle : les estimations varient entre 140 Gt (giga tonnes ou milliards de tonnes, équivalent à 1050 Gb ou giga barils) d'après le Oil and Gas Journal (OGJ) et 160 Gt (1200 Gb) d'après l'US Geological Survey (USGS). En se basant sur la consommation actuelle de pétrole, ces chiffres nous mèneraient entre 53 et 63 ans de production (suivant les ressources). Le tableau ci-dessous indique la répartition des réserves de pétrole prouvées en fonction des différentes zones géographiques de production.

### Part des régions dans les réserves mondiales prouvées de pétrole brut en pourcentage [int]



Le fait marquant est une disproportion énorme entre les réserves du Moyen Orient et celles du reste du monde. De plus, en distinguant une répartition non par région mais par pays de l'OPEP\* et hors OPEP, celle-ci montre que les premiers détiennent 78% des réserves mondiales de pétrole brut prouvées.

La consommation totale cumulée de pétrole jusqu'au début de ce millénaire atteint 110 Gt (825 Gb), soit moins que les réserves prouvées. Cette constatation amène certains géologues à prédire l'apparition de tensions dans le secteur du pétrole en se basant sur la méthode de King Hubert, géologue américain qui prédit, en 1956, le déclin de la production américaine pour 1970, soit l'année à partir de laquelle la moitié des réserves de pétrole eurent été exploitées, prédiction qui se confirma. Si ce modèle s'applique non plus seulement aux Etats-Unis mais à l'ensemble de la planète, une crise de la production pétrolière pourrait survenir dans un futur proche.

En 2002 les coûts de développement des bruts conventionnels se situent entre 2 et 15 dollars américains le baril et ces coûts diminuent en moyenne de 1 dollar américain par an. Ce faisant, certains bruts non-conventionnels deviennent exploitables. Ainsi, dans le passé, l'exploitation "off shore" était au début considérée comme beaucoup trop chère et presque impossible, alors qu'aujourd'hui plus personne ne remet en cause la rentabilité de ces explorations et pour autant le prix du brut en terme réel ne s'est pas envolé.

Les bruts non-conventionnels regroupent les nappes difficiles d'accès comme les nappes sous

couche de sel, "l'off shore" profond et très profond (la couche d'eau étant de plus de 1000 mètres voire de plus de 2000 mètres), les huiles extra lourdes, les sables asphaltiques et les schistes bitumineux.

Concernant "l'off shore" profond et très profond, des forages existent déjà qui permettent d'exploiter des gisements à plus de 1000 mètres de profondeur sous la surface de la mer. C'est le cas au Congo-Brazaville et en Angola (dont un gisement en mer profonde est le plus grand d'Afrique), mais aussi au Gabon avec un forage à plus de 2500 mètres.

La frontière entre brut non-conventionnel et brut conventionnel est sans cesse repoussé par les avancées technologiques. Ainsi les huiles extra-lourdes de l'Orénoque au Vénézuéla, n'étaient autrefois exploitables que pour un prix du baril supérieur à 30 dollars américains. La diminution du coût de développement rend aujourd'hui possible cette exploitation pour un prix du baril à 15 dollars américains.

Outre ces considérations sur les réserves de pétrole, en terme de longévité de la ressource pétrole, les progrès technologiques permettent de substituer le pétrole par d'autres produits notamment dans la fabrication de carburants liquides. A titre d'illustration, le procédé de Fischer Tropsh permet de fabriquer un carburant liquide pour véhicule à partir de gaz naturel ou même de pétroles non-conventionnels difficilement exploitables tels quels. Un autre procédé propose la liquéfaction du charbon afin d'obtenir également des carburants liquides. L'épuisement des réserves de pétrole est donc un sujet ouvert qu'il est très difficile, voire impossible, de prédire aujourd'hui. [int]

## **GLOSSAIRE**

\* : Les pays membres de l'Organisation des pays exportateurs de pétrole (OPEP ou OPEC en anglais) sont l'Algérie, l'Indonésie, l'Iran, l'Iraq, le Koweït, la Libye, le Nigéria, le Qatar, l'Arabie Saoudite, les Emirats Arabes Unis et le Vénézuéla.

Chaîne aliphatique : chaîne d'atomes de carbone linéaire qui peut être ramifiée. Les molécules aliphatiques sont grasses.

Hexane : hydrocarbure contenant six carbones. Eventuellement, la chaîne carbonée peut formée un cycle, c'est alors un cyclohexane.

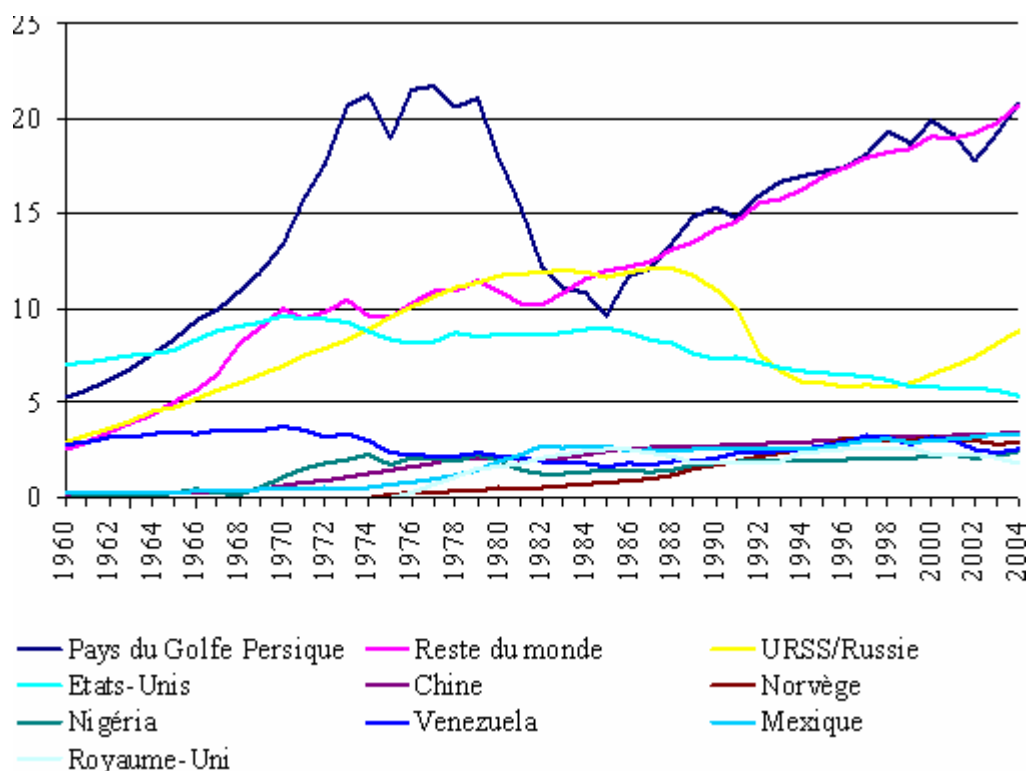
Pentane : hydrocarbure contenant cinq carbones. Eventuellement, la chaîne carbonée peut formée un cycle, c'est alors un cyclopentane.

Off shore : une exploitation pétrolière est dite "off shore" lorsqu'elle se trouve en pleine mer. L'exploitation se fait alors à partir de plates-formes flottantes ancrées au fond de la mer.

## I-4 Production :

Traditionnellement la majeure partie de la production est le fait d'un petit nombre de pays. La décennie des années 1970 a marqué le début de la production pour de nouveaux pays : Norvège, Royaume Uni, Mexique, Nigéria, Rép. dém. pop. de Chine, etc. Deux facteurs expliquent ce phénomène. Le premier est relatif à la crise pétrolière de 1973 qui s'est traduite par une hausse du prix du brut, rendant ainsi rentable des gisements qui ne l'étaient pas jusqu'alors (du fait de coûts technologiques élevés). Le deuxième tient plus au développement économique de certains pays (c'est par exemple le cas de la Rép. dém. pop. de Chine) et/ou à l'importance des investissements étrangers notamment dans le secteur pétrolier (c'est par exemple le cas du Nigéria).

**Production de pétrole brut par principaux pays ou groupes de pays depuis 1960, en millions de barils par jour [Int]**



Les pays du Golfe Persique, principaux producteurs de pétrole avant la double crise des années 1970 n'ont toujours pas retrouvé leur niveau de production d'alors, et ce bien que la consommation et la production mondiale de pétrole aient triplé de 1960 à 2004. La concurrence de nouveaux producteurs et les politiques de soutien des prix de la part de l'OPEP par une limitation de leur production expliquent en grande partie cette position des pays du Golfe. [1], [Int]

<b>La production de pétrole dans le monde en 1973 et 2000 en million de tonnes</b>				
	<b>1973</b>		<b>2000 p</b>	
		<b>%</b>		<b>%</b>
<b>Amérique du Nord</b>	<b>635</b>	<b>22,2</b>	<b>656</b>	<b>18,4</b>
Canada	94	3,3	126	3,5
Etats-Unis	513	18,0	359	10,1
Mexique	27	1,0	171	4,8
<b>Amérique latine</b>	<b>245</b>	<b>8,6</b>	<b>350</b>	<b>9,8</b>
dont : Brésil	8	0,3	64	1,8
Vénézuéla (1)	178	6,2	172	4,8
<b>Europe de l'Ouest (2)</b>	<b>23</b>	<b>0,8</b>	<b>323</b>	<b>9,1</b>
dont : Norvège	2	0,1	158	4,4
Royaume-Uni	1	0,0	127	3,6
<b>Europe de l'Est (3)</b>	<b>449</b>	<b>15,7</b>	<b>380</b>	<b>10,5</b>
dont : ex-URSS	429	15,0	393	11,0
<b>Afrique</b>	<b>289</b>	<b>10,1</b>	<b>366</b>	<b>10,3</b>
dont : Algérie (1)	51	1,8	66	1,9
Angola	8	0,3	37	1,0
Egypte	9	0,3	39	1,1
Libye (1)	106	3,7	68	1,9
Nigeria (1)	101	3,5	103	2,9
<b>Moyen-Orient</b>	<b>1 054</b>	<b>36,9</b>	<b>1 089</b>	<b>30,6</b>
dont : Arabie Saoudite (1)	380	13,3	427	12,0
Emirats Arabes Unis (1)	74	2,6	112	3,2
Iran (1)	293	10,3	186	5,2
Irak (1)	100	3,5	127	3,6
Koweït (1)	151	5,3	102	2,9
Oman	15	0,5	49	1,4
<b>Extrême-Orient</b>	<b>145</b>	<b>5,1</b>	<b>340</b>	<b>9,5</b>
dont : Chine	54	1,9	163	4,6
Indonésie (1)	66	2,3	65	1,8
<b>Océanie</b>	<b>19</b>	<b>0,7</b>	<b>34</b>	<b>0,9</b>
<b>Monde</b>	<b>2859</b>	<b>100,0</b>	<b>3450</b>	<b>100,0</b>
dont : OPEP	1 528	53,4	1465	41,1
non OPEP	1 331	46,6	2 096	58,9
non OPEP hors ex-URSS	902	31,6	1 703	47,8

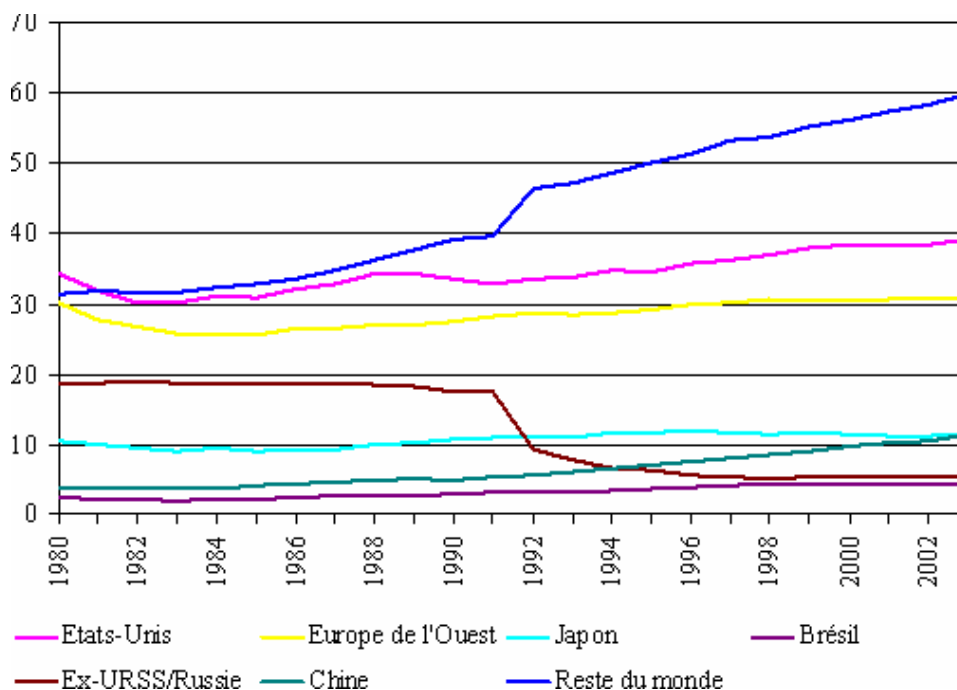
p : provisoire  
1 : Pays membres de l'OPEP, voir Annexes  
2 : Europe OCDE, y compris la Hongrie, la Pologne et la République Tchèque et la Slovaquie  
3 : Europe non OCDE



## I-5 Consommation :

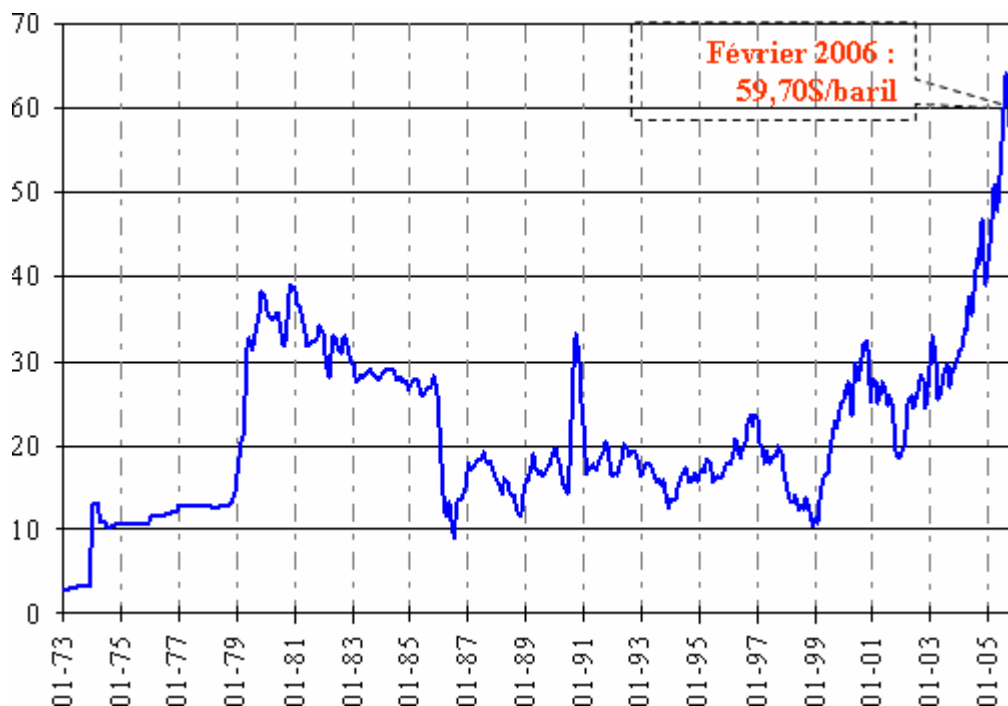
Le pétrole étant la principale source d'énergie dans le monde, il n'est pas surprenant de retrouver les mêmes pays comme principaux consommateurs d'énergie et de pétrole. Les crises de 1973 et 1979 ont eu comme conséquence de réduire la dépendance énergétique par rapport au pétrole, en particulier en Europe. Un changement de politique énergétique est intervenu et il s'est orienté vers deux objectifs principaux, d'une part la réduction de la facture énergétique de manière globale (grâce à l'application de nouvelles normes d'isolation des bâtiments, à l'amélioration du rendement des moteurs à combustion, à une taxation plus importante, etc) et d'autre part la promotion de sources d'énergies autres que le pétrole (avec, par exemple, la construction de centrales électriques nucléaires). Ce changement fut très important en Europe, et ce n'est qu'à partir de 1986, après sept années consécutives de baisse, qu'une hausse de la consommation de pétrole apparaît de nouveau. De plus, cette augmentation est progressive et en 2003 le niveau de consommation de pétrole européen n'a pas rattrapé celui de 1979. Aux Etats-Unis la consommation de pétrole repart très tôt à la hausse après les crises, dès 1983. Ce n'est cependant qu'en 1999 que la consommation a retrouvé son niveau de 1979. Les politiques énergétiques des autres pays furent moins marquées par ces crises parce qu'ils ne dépendaient pas autant du pétrole ou que leur production propre suffisait à alimenter leurs besoins (c'est le cas pour la Rép. Dém. pop. De Chine et l'ex-URSS). [1], [2]

**Consommation de pétrole brut en millions de milliards de BTU de 1971 à 2003 [Int]**



## I-6 L'évolution des prix du pétrole :

Prix nominaux d'une moyenne des prix du Brent FOB UK, du WTI FOB USA Gulf et du DUBAÏ FOB Dubaï 1973 - Février 2006 (en dollars par baril) [Int]



## I-7 Secteurs d'utilisation :

Le pétrole est un produit stratégique utilisé dans un grand nombre de secteurs différents, ce qui en fait une denrée vitale et centrale dans l'économie mondiale. Un baril de pétrole contient 42 gallons US, soit 158,97 litres (1 gallonUS = 3,785 litres). SA titre d'exemple, en moyenne à la sortie d'une raffinerie aux Etats-Unis, un baril de pétrole génère les quantités suivantes de produits raffinés

Produit	Litres
carburants	73,8
gazole et mazout léger (fioul domestique)	34,8
kérosène (carburacteur pour l'aviation)	15,2
mazout lourd (fret maritime, centrales thermiques, industries)	8,7
gaz de pétrole du raffinage	7,2

autres gaz (éthane, propane, butane)	7.2
coke	6.8
asphalte	4.9
bases pour la pétrochimie	4.5
lubrifiants	1.9
kérosène (ou pétrole lampant)	0.7
autres (cires, graisses)	1.1

Ces chiffres varient d'un type de pétrole à un autre mais aussi d'une raffinerie à une autre et d'une époque à une autre. En effet, il est possible à partir du même brut de raffiner plus d'essence (en augmentant la capacité de crackage catalytique) ou plus de fioul en modifiant les procédés utilisés.

Ci-dessous, un éventail non exhaustif des différents secteurs par produit. [3]

<b>Les carburants (GPL, essence, gazole, carburéacteur)</b>	Industrie automobile, aviation
<b>Les combustibles (fioul domestique, fioul lourd)</b>	Chauffage domestique, chauffage urbain, transport maritime, locomotives, fours industriels. Dans l'industrie : cimenteries, tuileries, papeteries, verrerie et usines thermiques (production d'électricité).
<b>Les lubrifiants</b>	Huiles pour moteur, huiles pour transmission automatique ou hydraulique. Huiles de procédés : utilisées dans les encres, les insecticides, le caoutchouc Huiles blanches pharmaceutiques. Graisses : essentiellement dans les roulements à bille. Les cires : encaustiques, enduction de tissu, protection des métaux.

	Les paraffines : emballages, produits d'entretien, produits pharmaceutiques, explosifs, imperméabilisation du bois.
<b>Les bitumes</b>	Revêtement routier, étanchéité des bâtiments, isolation, anticorrosion...
<b>Les gaz</b>	Usage domestique
<b>Les bases pétrochimiques</b>	Polymères : plastique, isolants. Fibres synthétiques : nylon. Caoutchoucs synthétiques. Solvants : encres d'imprimerie, peinture, colles,

## Chapitre II :

# Etude des propriétés physiques et thermodynamiques

## II-1 Introduction :

Comme il a été déjà évoqué, le premier traitement subi par le pétrole brut en raffinerie est une distillation sous pression atmosphérique qui classe les hydrocarbures en fonction de leur point d'ébullition et permet ainsi d'obtenir une série de coupes ou fractions : gaz, essence légère, essence lourde, kérosène, gasoil et résidu. Ces produits sont sélectionnés pour satisfaire essentiellement les spécifications de distillation. Les caractéristiques des coupes obtenues par ce fonctionnement, ou bien satisfont les autres spécifications et le produit peut être considéré comme commercial, ou bien ne les satisfont pas et la fraction n'est qu'un produit intermédiaire qui devra subir des transformations ultérieures qui, elles même, exigent souvent des spécifications de composition. Quelle que soit la destination de ces coupes, les analyses du pétrole brut et leurs traductions graphiques permettent d'en estimer directement les rendements avec une excellente précision, c'est-à-dire que la distillation T.B.P du laboratoire fournit des renseignements qui sont extrapolables aux unités industrielles.

La complexité de la composition des fractions pétrolières n'en permet pas l'analyse, il est difficile de connaître les propriétés physiques et thermodynamiques, si ce n'est certaines fois par des méthodes expérimentales élaborées. Seules seront connues, en général, la densité et la courbe de distillation A.S.T.M. qui permettent par l'intermédiaire du facteur de caractérisation, de se faire une idée de la structure moléculaire moyenne. Dans ce domaine la règle générale consiste à trouver une équivalence entre la fraction complexe et un hydrocarbure pur fictif qui aurait les mêmes caractéristiques physico-chimiques que la fraction. Cette équivalence portera, soit sur le poids moléculaire, soit sur la température d'ébullition, soit sur les coordonnées critiques, etc.

Cependant, il devient de plus en plus nécessaire pour les raffineurs ou le chercheur de disposer rapidement des propriétés physiques des produits pétrolières.

De nombreuses études ont été menées dans le but de proposer des méthodes pour la prédiction des paramètres physiques et thermodynamiques des fractions pétrolières. Ces travaux visent à minimiser le temps d'accès à ces paramètres et éviter autant que faire ce peu les manipulations lentes, fastidieuses et coûteuses.

Dans le cadre de ce chapitre, nous avons décrit brièvement quelques méthodes de calcul existantes dans la littérature.

## II-2 Les propriétés critiques :

(Critique) ce terme a été proposé par **Andrews (1869, 1876, 1887)** dans ses travaux sur les propriétés du dioxyde de carbone. Les travaux de **Van der Waals (1873)** et développement de la loi des états correspondants furent le point de départ d'une multitude de méthodes expérimentales visant à mesurer les propriétés critiques des corps purs et des mélanges.

La connaissance des propriétés critiques est indispensable dans le traitement des mélanges d'hydrocarbures où la coexistence d'une phase gazeuse et d'une phase liquide est fréquente. Ils permettent de prévoir le comportement des fluides dans un large intervalle de température et de pression et définir ainsi les conditions d'existence de leurs différentes phases. Ce sont également les paramètres clés de la prédiction des propriétés thermodynamiques et volumétriques des corps purs et des mélanges de corps purs par le biais de la loi des états correspondants.

Dans un domaine de coordonnées pression- température, la pression et la température critiques sont les coordonnées du point critique qui représente la borne supérieure au dessus de laquelle la coexistence des deux phases liquide et vapeur n'est plus possible. Il correspond donc à la limite supérieure de la courbe de tension de vapeur, au-delà de laquelle le changement de phase n'est plus observé. Dans le cas d'un mélange, il se trouve à la jonction des courbes de bulle et de rosée.

Le point critique correspond à l'identité parfaite des propriétés du liquide et de la vapeur. La séparation d'un tel mélange est donc impossible en ce point où la chaleur de vaporisation est nulle. [3], [9].

### II-2-1 Mesure expérimentale des paramètres critiques :

Le principe de la mesure consiste à chauffer un fluide jusqu'à la disparition du ménisque séparant les phase liquide et vapeur et apparition d'un brouillard opalescent. La température, la pression et le volume du fluide, à cet instant précis, correspondent aux propriétés critiques de la substance étudiée. Suite aux travaux de **De la Tour**, plusieurs appareils de mesure des paramètres critiques ont été décrits par un grand nombre d'auteurs tels que **Andrews (1869, 1876, 1887)**, **Ramsay et Young (1887)**, **Cailletet (1869)**, **Ambrose et Toxsend (1963)**, **Quadri et al. (1991)** et bien d'autres auteurs. Les équipements modernes ne diffèrent pas de point de vue principe des premiers appareils conçus pour effectuer cette mesure. Les équipements utilisés de nos jours sont généralement dotés dispositifs de contrôle, de régulation, d'acquisition, de visualisation automatique et pilotés par ordinateurs dotés de

logiciels spécifiques. Les contraintes environnementales de plus en plus sévères ont conduit les concepteurs d'appareils modernes de mesure de ces paramètres à éliminer le mercure utilisé comme fluide de déplacement.

Les valeurs expérimentales des propriétés critiques des corps purs ont été compilées par différents auteurs dans des banques de données dont les plus importantes sont celles établies par **Ambrose (1980)**, **Kudchadker, Alani et Zwolinsky (1968)** pour les composés organiques et **Mathews (1972)** pour les composés inorganiques et ce, en plus banques du **DIPPR**, de l'**API** et de **TRC**. [3], [10].

## II-2-2 Estimations des propriétés critiques des fractions pétrolières :

Les fractions pétrolières sont des mélanges très complexes, dont l'analyse quantitative et qualitative est difficile et onéreuse. Des corrélations ont été proposées afin d'estimer les propriétés critiques des fractions pétrolières.

Si la composition a pu être déterminée par des méthodes physiques d'analyse, pour les coupes légères les lois d'additivité des mélanges peuvent être utilisées. Mais ne donnent pas toujours une bonne précision pour toutes les propriétés

### II-2-2-1 Température critique des fractions pétrolières [3],[7] :

Les températures critiques des fractions pétrolières peuvent être déterminées en utilisant l'additivité lorsque la composition par familles chimiques est disponible. Dans le cas contraire, de nombreuses corrélations qui nécessitent la connaissance d'autres paramètres physiques tels que la température moyenne d'ébullition et la specific-gravity (densité) sont utilisées. Nous citons quelques unes des méthodes parmi les plus connues.

- **Lee-Kesler :**

$$T_{pc} (^{\circ}R) = 341.7 + 811 \times SpGr + (0.4244 + 0.1174 \times SpGr) \times T_{mav} + (0.4669 - 3.2623 \times SpGr) \times 10^5 \div T_{mav}.$$

$T_{pc}$  : Température pseudo critique en ( $^{\circ}R$ ).

$T_{mav}$  : Température moyenne d'ébullition en ( $^{\circ}R$ ).

$SpGr$  : specific gravity.

- **Riazi et Daubert (1980) :**

$$T_{PC} (^{\circ}R) = 24.2787 \times T_{mav}^{0.58848} \times SpGr^{0.3596}.$$

Avec  $T_{mav}$  : Température moyenne d'ébullition (temperature mean average) en  $^{\circ}R$



- **Mathur :**

$$T_{pc} (K) = 87.5 * M^{0.4406}.$$

$M$  : Masse molaire.

$T_{pc}$  : Température pseudo critique en kelvin

- **Corrélation de Penn State :**

$$T_{pc} (°F) = e^{(3.9935 \times T_{mav}^{0.8615} \times SpGr^{0.4614})}$$

- **Corrélation de Souahi et al. :**

Cette corrélation est basée sur la méthode de contribution des groupes.

$$c = a + \frac{b}{1 + \left(\frac{N}{c}\right)^d}$$

$N$  : le nombre d'atome de carbone.

$\theta_c$  : Paramètre critique.

$a, b, c$  et sont des constantes tabulées.

### II-2-2-2 Pression critique des fractions pétrolières [3],[7],[12] :

Pour l'estimation de la pression critique des fractions pétrolières, de nombreuses corrélations ont été établies par différents auteurs. Parmi les plus utilisées nous citons :

- **Riazi et Daubert (1980) :**

$$P_{pc} (psia) = 3.12281 \times 10^9 * T_{mav}^{-2.3125} * SpGr^{2.3201}.$$

$P_{pc}$  : Pression pseudo critique en psia.

- **Penn-State :**

$$P_{pc} (psia) = 3.4824 \times 10^9 \times T_{mav}^{-2.3177} \times SpGr^{2.4853}.$$

- **Corrélation de Lee-Kesler :**

$$\ln(P_{pc}) = 8.363 \times 566 \div SpGr \times (0.24244 + 2.29 \div SpGr + 0.11857 \div SpGr^2) \times 10^3 \times T_{mav} \\ + (1.4685 + 3.648 \div SpGr + 1.6977 \div SpGr^2) \times 10^{-10} \times T_{mav}.$$

$P_{pc}$  : Pression pseudo critique en psia

$T_{mav}$  : Température moyenne d'ébullition en °R

- **Corrélation de Mathur :**

$$P_{pc} (atm) = 4532 \times M^{-0.979}.$$

- **Corrélation de Souahi et al**

Cette corrélation est basée sur la méthode de contribution des groupes.

$$\theta_c = a + \frac{b}{1 + \left(\frac{N}{c}\right)^d}$$

N : le nombre d'atome de carbone.

$\theta_c$  : Paramètre critique.

### II-2-2-3 Volume critique des fractions pétrolières [3],[7] :

Plusieurs auteurs ont proposé des corrélations empiriques pour l'estimation du volume critique des fractions pétrolières de même type que celles proposées pour les températures critiques.

Parmi les corrélations les plus utilisées nous citons :

- **Riazi et Daubert (1980) :**

$$V_{pc} (ft / lb^3) = 7.5214 \times 10^5 \times T_{mav}^{0.2896} \times SpGr^{0.7666} . \mathbb{K}$$

$V_{pc}$  : Volume pseudo critique en  $ft / lb^3$

### II-3 Facteur acentrique [3], [12], [7] :

Le facteur acentrique  $\omega$  a été proposé par **Pitzer (1955)** pour caractériser la non-sphéricité des molécules, de telle façon que ce facteur s'annule pour les molécules sphériques les plus simples telles que l'argon, le xénon et le krypton. Il a été défini, conventionnellement, à la température réduite de **0.7** sur la courbe d'équilibre liquide-vapeur par l'équation suivante :

$$\omega = -\log(P_{vap} \div P_c)_{T_r=0.7} - 1$$

Où  $P_{vap}$  est la tension de vapeur et  $P_c$  la pression critique, exprimées toutes deux en atmosphères.

Ce facteur a joué un rôle très important dans l'élaboration des corrélations des états correspondants à trois paramètres et dans le développement des équations d'état empiriques.

Les facteurs acentriques des séries homologues augmentent avec le nombre d'atomes de carbone. Les valeurs les plus faibles sont obtenues dans le cas des hydrocarbures naphténiques et les plus élevées pour les hydrocarbures paraffiniques. Les aromatiques possèdent quant à eux, des facteurs intermédiaires.

Lorsque les valeurs expérimentales des pressions de saturation et des pressions critiques ne sont pas disponibles, nous pouvons estimer les facteurs acentriques à partir des corrélations proposées par différents auteurs.

**Edmister (1958)** est le premier à avoir proposé une méthode d'estimation du facteur acentrique en se basant sur l'équation de tension de vapeur d'**Antoine**. Cette équation est valable aussi bien pour les corps purs que pour les fractions pétrolières. Elle est donnée par l'expression suivante :

$$\omega = \frac{3}{7} \times \left[ \frac{\log P_c}{\frac{T_c}{T_{mav}} - 1} \right] - 1$$

Lee et Kesler (1975) ont développé une corrélation pour l'estimation des facteurs acentrique des substances à partir de la connaissance de leur température d'ébullition et des températures et pression critiques. Ces mêmes auteurs ont établi en 1976 une

équation pour la prédiction de ce facteur dans le cas des fractions pétrolières lourdes dont les températures réduites sont supérieures à **0.8**.

Plus tard, **Watanasiri, Owens et Starling (1985)** ont également proposé une équation empirique pour l'estimation du facteur acentrique des hydrocarbures et fractions pétrolières à partir de la masse molaire, la température d'ébullition et la specific-gravity.

Par ailleurs, **Riazi et EL-Sahhaf (1995)** ont proposé une équation de calcul des facteurs acentriques des hydrocarbures de type n-alcanes, n-alkylcyclopentanes, n-alkylcyclohexanes et n-alkylbenzènes à partir de la connaissance de la masse molaire uniquement.

**Skander, Souahi et Chitour (1999)** ont eux également élaboré des équations spécifiques à des familles d'hydrocarbures. Leur corrélation est une équation de forme polynomiale de type :

$$\omega = a + b \times N_c + c \times N_c^2 + d \times N_c^3 + e * \times N_c^4$$

$N_c$  : Nombre d'atomes de carbone

Les constantes a, b, c, d et e dépendent de la famille d'hydrocarbure.

# Chapitre III

## Les algorithmes génétiques

### **III-1 Introduction :**

Parmi les méthodes stochastiques les plus utilisées à l'heure actuelle, les algorithmes génétiques suscitent un intérêt croissant chez les chercheurs et ce, depuis une trentaine d'années. Ces algorithmes ont d'abord été développés par Holland [HOLL-1975].

L'idée de base est que les membres les plus performants d'une population d'individus dans un environnement donné ont une probabilité supérieure de voir leur information génétique servir à la génération suivante, c'est la sélection naturelle vue par Darwin. La transmission du patrimoine génétique est effectuée par le biais de reproductions (sélections et croisements), auxquels viennent se greffer des mutations du code génétique qui apparaissent aléatoirement.

La traduction mathématique est très simple à programmer. Plusieurs individus (une population) sont définis en réalisant des analogies avec des variables réelles (leur génome correspond aux variables commandables  $u$ ). Leur capacité à évoluer dans leur environnement est traduite par l'évaluation d'une fonction objectif  $F$  appliquée à leur génome. Une fois que deux individus ont été sélectionnés, ils sont croisés (échange d'une partie de leur génome). Deux enfants sont ainsi créés. En effet, dans l'application présente et à cause du grand nombre de générations, on a fait le choix de travailler à population constante. Deux autres individus (parmi les plus mauvais) sont donc sélectionnés et éliminés.

Bien que, pour les comprendre réellement, les fondements et les descriptions mathématiques de ces algorithmes soient encore à développer, ils se révèlent souvent très efficaces et sont de plus en plus utilisés dans tous les domaines de l'optimisation. Par ailleurs, leur programmation est très simple.

#### **III-1-1 Définition d'un individu :**

Un individu est une configuration possible des variables d'optimisation. En génie chimique, un individu correspond donc à une solution possible de paramètres de conception ou optimum d'un coefficient de modèle, de configuration ou de réglage des divers éléments. Si l'on considère  $n$  variables  $x_i$ , un individu sera l'ensemble :  $\{x_1, x_2, \dots, x_n\}$  représentant une configuration possible de ces variables.

### **III-1-2 Codage des variables :**

Afin de développer l'analogie avec la génétique, les variables sont codées le plus souvent sous forme binaire. Elles apparaissent alors sous la forme d'une chaîne de 0 et de 1 appelés **gènes**.

Une variable étant codée sous forme binaire, un individu est représenté par la concaténation de ces chaînes. Cette concaténation est appelée **chromosome**.

La puissance et la généralité de ce codage proviennent du fait que l'on peut aussi bien manipuler des entités numériques (équivalent binaire/réel par une table de conversion) que des objets (en modélisant leur présence ou absence).

D'autres codages peuvent être aussi utilisés. On a utilisé un codage d'entiers servant de pointeurs dans des tables ou données constructeurs. C'est ainsi que peut être intégré le rapport de transformation discret du régleur en charge.

De manière générale, l'association d'un codage particulier (le génotype) à un phénomène donné (le phénotype) est un facteur déterminant pour l'efficacité de l'algorithme. Ainsi, le codage est adapté en fonction du problème à traiter.

### **III-1-3 Définition d'une population :**

Un ensemble d'individus constitue une population. Au sein de la population, chaque individu est caractérisé par sa performance souvent notée fitness. Cette dernière est égale à la fonction objectif lorsque le problème n'est pas contraint ou à la fonction objectif pénalisée lorsque les contraintes sont prises en compte. Le but des algorithmes génétiques est de trouver un optimum (i.e. un individu avec la meilleure performance possible). Un brassage des caractères de la population est donc effectué afin de permettre à des individus de plus grande performance d'émerger par croisement ou élitisme. Les croisements de ces éléments (enfants) seront les éléments d'une nouvelle population qui elle-même permettra de créer de nouvelles générations.

La création des enfants passe par trois opérations de base :

- La reproduction / sélection
- Les croisements
- La mutation

### III-2 Déroulement de l'algorithme :

La figure précédente présente l'organigramme de la programmation des algorithmes génétiques qui sont utilisés à population constante. On a rajouté la mutation déterministe pour tenir compte du caractère mixte du problème de caractérisation des fractions pétrolières

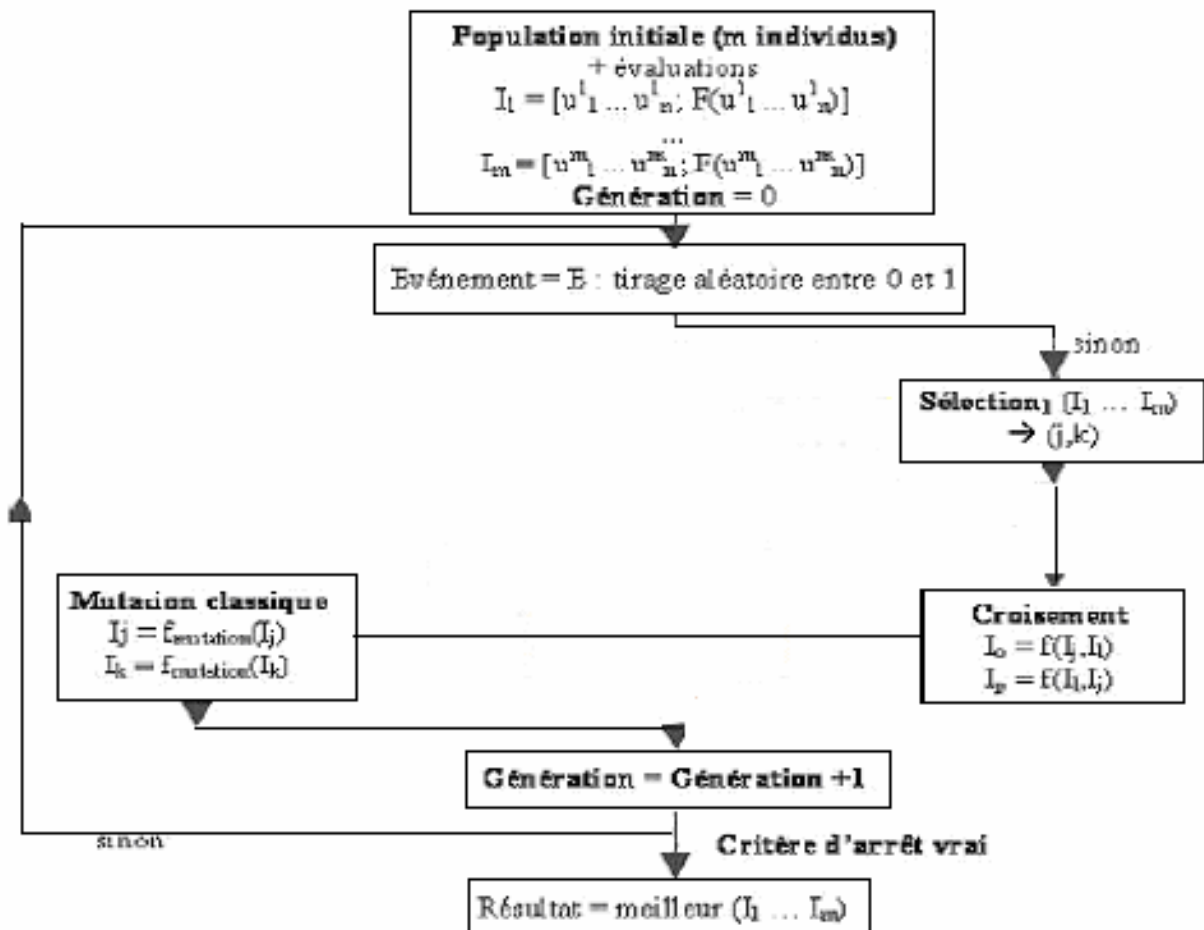


Figure III – 1 : Organigramme des algorithmes génétiques

Les notations utilisées dans l'organigramme précédent sont :

- $I_k$  est le  $k^{\text{ième}}$  individu de la population totale,
- $u_1^k \dots u_n^k$  sont les valeurs des  $n$  variables de commande de l'individu  $k$ ,



- $F(u_1^1 \dots u_n^1)$  est l'évaluation de la fonction objectif  $F$  appliquée à l'individu 1,
- Événement =  $E$  est un nombre aléatoire entre 0 et 1, il définit l'événement qui va se produire (mutation si  $E$  est inférieur à un taux donné  $T_{\text{mutation}}$  (%) ou croisement sinon),
- pour « distance  $\leq$  Taux », l'opérateur distance détermine si trop de points (Taux) sont trop proches (norme sur les distances à définir voir annexe) il y a alors mutation forcée (prioritaire par rapport à l'événement  $E$ ).
- Sélection<sub>1</sub> est l'opérateur sélection qui permet de choisir deux individus souvent parmi les meilleurs,
- Sélection<sub>2</sub> est l'opération qui permet de choisir deux individus parmi les plus mauvais et les éliminer pour maintenir la population constante,
- $f$  est l'opérateur de croisement (uniforme ou non-uniforme),
- $f_{\text{mutation}}$  est l'opérateur de mutation,
- Critère d'arrêt est un opérateur qui donne la fin du processus d'optimisation (il peut être lié à une interpolation de type exponentielle de la courbe de décroissance ou être lié à un nombre de génération fixé).

### III-2-1 Les opérations :

#### a- La Reproduction / Sélection :

Afin de créer une nouvelle population, il est important de trouver un critère pour sélectionner les individus qui serviront de parents pour les futurs individus. C'est à ce niveau que rentre en compte le principe Darwinien de survie du plus performant. En effet, toutes les méthodes de sélection sont basées sur la classification des individus suivant leurs performances respectives. Elles sont construites de façon à donner le plus de chance aux individus les plus performants. Trois méthodes sont communément utilisées : l'élitisme, le tournoi et la roue biaisée.

##### a- 1 L'élitisme :

Deux parents sont choisis aléatoirement parmi les meilleurs de la population. Un taux d'élitisme est souvent utilisé, il correspond aux  $n$  meilleurs parmi la population toute entière.

**a- 2 Le tournoi :**

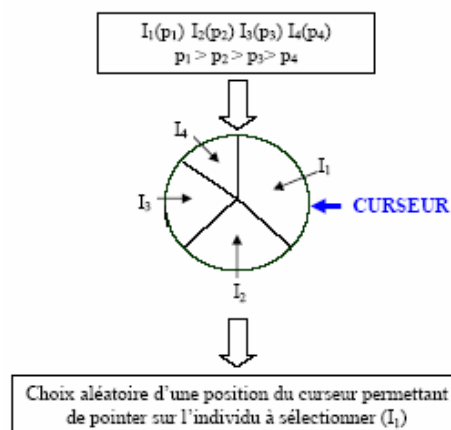
Deux parents sont choisis de manière aléatoire. Leurs performances respectives sont ensuite comparées : seul le parent le plus performant est retenu et apporte sa contribution à la génération suivante. Cette opération est répétée deux fois afin d'obtenir deux parents. Leur contribution à la prochaine génération sera opérée par leur croisement, défini par la suite.

**a-3 La roue biaisée :**

Afin de représenter la performance d'un individu par rapport aux autres membres de la population, la performance totale d'une population est représentée par une roue complète. Cette roue est divisée en autant de secteurs que d'individus. Plus la performance d'un individu est importante plus la taille du secteur correspondant est importante.

Une fois la roue définie, un nombre aléatoire est tiré décrivant la position d'un curseur autour de la roue. Lorsque le curseur tombe sur le secteur correspondant à un individu précis, ce dernier est sélectionné.

Il apparaît clairement que ce processus favorise les individus de performance élevée sans pour autant les choisir systématiquement et dédaigner les autres. Le ratio de taille des secteurs peut être paramétré entièrement, ce qui permet d'avoir un élitisme important (grosse différence) ou un élitisme faible (même probabilité pour tous les individus). On a choisi de donner une valeur exponentielle décroissante aux secteurs pour réaliser la sélection par roue biaisée.



**Figure III 2 : Sélection d'un individu par la roue biaisée individus distincts**

Une fois deux parents sélectionnés, ils seront croisés, c'est-à-dire que certaines de leurs caractéristiques vont être échangées pour obtenir un enfant. Différents types de croisement peuvent alors être utilisés.

## b. Les croisements :

### b-1 Le croisement uniforme :

A partir de deux parents un enfant est créé, constitué de gènes provenant aléatoirement de l'un ou l'autre des parents et ce sur la totalité du chromosome. Pour sélectionner le parent qui fournira un gène donné, un nombre aléatoire compris entre 0 et 1 est tiré: en fonction de sa valeur par rapport à 0.5, l'enfant héritera du gène de l'un ou de l'autre des parents (exemple figure suivante). Un deuxième enfant est aussi créé.

### b-2 Le croisement à une ou deux coupures :

Après avoir été choisi, chaque parent va être scindé en deux ou trois parties puis les deux parents vont échanger entre eux certaines de ces parties. Dans le cas d'un croisement à une coupure, une position est choisie aléatoirement sur les chromosomes qui échangeront leurs deux extrémités.

Les croisements à 2 coupures divisent les chromosomes en 3 parties (figure suivante). L'enfant sera le résultat de l'échange des parties centrales. Les échanges se font suivant une certaine probabilité  $\alpha$ , grâce au tirage d'un nombre aléatoire compris entre 0 et 1. Si ce nombre est supérieur à  $\alpha$ , l'échange a bien lieu.

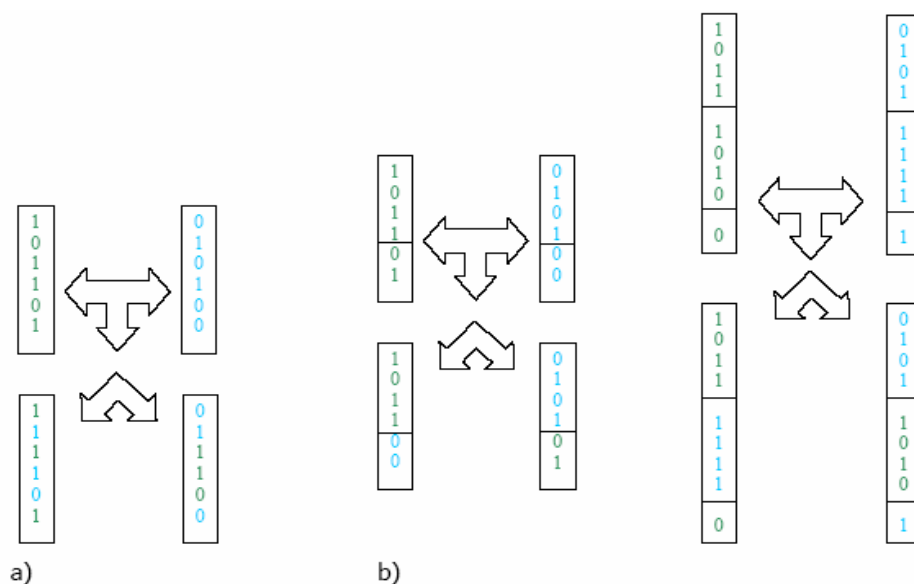
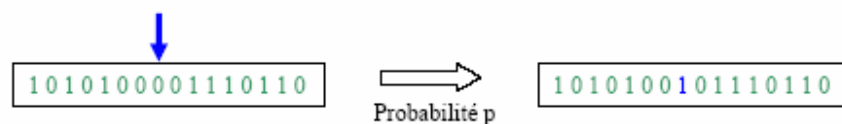


Figure III - 3 : Croisement uniforme (a) et par une ou deux coupures (b)

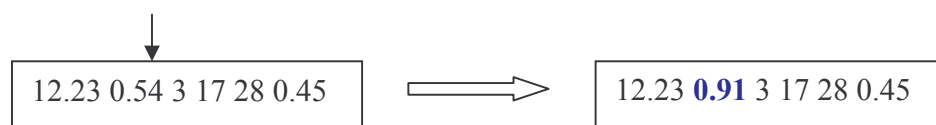
### c- La mutation :

La mutation est un opérateur qui permet de changer la valeur d'un gène, choisi aléatoirement, et ce avec une certaine probabilité. Si le  $i$ ème individu est sélectionné pour une mutation alors la valeur d'un de ses gènes sera transformée suivant une probabilité  $p$ . Cette transformation correspond le plus souvent au passage de 0 à 1 ou de 1 à 0 pour un codage binaire et un tirage aléatoire d'une valeur continue dans un intervalle donné pour un codage réel.

La mutation joue un rôle primordial dans la convergence d'un algorithme génétique. En effet au cours des itérations, les populations tendent à s'homogénéiser, c'est-à-dire à rassembler des individus très similaires. Ceci a pour effet de rapidement engager l'algorithme dans une direction de recherche bien précise et de le faire échouer sur un optimum local. Nous observons alors un appauvrissement du génome de la population complète et une parfaite incapacité à explorer la totalité de l'espace de recherche. La mutation permet de changer brusquement certains individus et ainsi d'explorer de nouvelles directions dans l'espace de recherche et ceci indépendamment du processus « déterminé » de sélection et de croisement.



**Figure III - 4 : Mutation aléatoire d'un des gènes codé sous forme binaire**



**Figure III - 5 : Mutation aléatoire d'un des gènes codé sous forme réelle**

Afin de traiter au mieux les problèmes d'optimisation hybride, une mutation « intelligente » a été mise au point. Lorsque l'événement mutation intelligente apparaît, un individu, qui peut être choisi parmi les meilleurs, subit une optimisation déterministe (voir optimisation linéaire sous contrainte) de ses gènes qui correspondent à un codage réel et une évolution continue de la fonction objectif. Ceci permet d'accélérer considérablement la recherche des optimums locaux,

### III-3 Prise en compte des contraintes par pénalisation :

Les algorithmes génétiques utilisant uniquement une quantité appelée fitness (ou évaluation numérique de la fonction objectif) pour sélectionner les individus, l'état de violation ou de Satisfaction des contraintes doit être inclus dans cette quantité. Toutes les méthodes de pris en compte des contraintes sont basées sur la pénalisation de la fonction objectif.

Soit une fonction objectif  $f(\mathbf{X})$  à minimiser sous un ensemble de  $m$  contraintes  $g_i(\mathbf{X}) \leq 0$ .

La performance globale  $P(\mathbf{X})$  est construite de la façon suivante :

$$P(\mathbf{X}) = f(\mathbf{X}) + q(\mathbf{X})$$

Où  $q$  est une fonction de la violation des contraintes. Elle peut prendre plusieurs formes :

- Pénalisation proportionnelle ou quadratique à la violation :

$$q(\mathbf{X}) = \sum p_i \max \{0, g_i(\mathbf{X})\}$$

Où  $p_i$  : coefficient de pénalisation à la  $i^{\text{ème}}$  contrainte

Les valeurs attribuées aux différents coefficients de contraintes  $p_i$  doivent souligner la violation. Pourtant des valeurs trop importantes des coefficients de pénalisation peuvent empêcher l'exploration d'une partie du domaine de recherche. Néanmoins, la puissance des Algorithmes Génétiques réside principalement dans la collecte d'informations sur l'ensemble du domaine, réalisable et non réalisable. Il existe donc une valeur optimale de ces coefficients à trouver, autre apport de l'expertise du programmeur.

Il semble que le choix optimal des coefficients de pénalisation par les Algorithmes Génétiques eux-mêmes soit une excellente alternative. Les coefficients sont alors inclus dans les chromosomes au même titre que les autres variables. Cette méthode est idéale pour des problèmes d'optimisation pour lesquels l'évaluation de la performance est très rapide, autorisant ainsi des tailles de population importantes.

### **III-4 Convergence et inconvénients des algorithmes génétiques :**

#### **4-1. Inconvénients :**

La convergence des algorithmes génétiques dépend fortement des paramètres suivants :

- La taille de la population. Si une trop petite taille ne permet pas d'échantillonner suffisamment l'espace de recherche pour assurer un optimum global, une taille trop importante peut rendre difficile la convergence sur un nombre raisonnable de générations.
- Le type de croisement. La littérature rapporte des résultats satisfaisants sur l'utilisation du croisement uniforme, ou du croisement à une ou deux coupures. Pourtant il faut souligner que le croisement uniforme n'est préférable que dans le cas de tailles de populations faibles, car il permet de maintenir une bonne diversité et d'éviter une convergence prématurée.
- Le taux de mutation. Un taux de mutation trop important peut ralentir la convergence de la population. En effet, la modification presque systématique d'un bit sur un chromosome diminue l'efficacité de la sélection et des croisements.

Les algorithmes génétiques nécessitent un réglage de ces paramètres, ce qui ne peut être effectué que par essais successifs. Cette lourdeur vient se rajouter aux temps de calculs nécessaires à l'évaluation des performances dans chaque population.

Pourtant les algorithmes génétiques restent un outil très puissant pour la gestion de problèmes à fonctions discontinues ou à variables discrètes. Le problème des temps de calculs prohibitifs des algorithmes génétiques peut aussi être surmonté par l'utilisation du calcul parallèle qui redistribue les calculs de performances d'une population entre plusieurs ordinateurs. La valeur de la performance de chaque individu est en effet un calcul indépendant des autres donc pouvant être distribué sans aucun problème.

#### **4-2. Critère de convergence :**

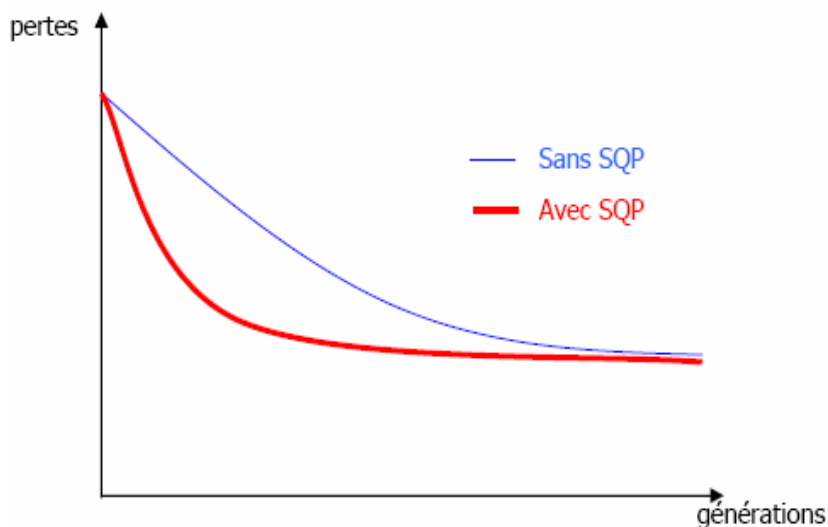
Il n'existe malheureusement pas de critères d'arrêt de type déterministe pour des méthodes stochastiques.

Certains auteurs utilisent souvent des critères basés sur leur expertise (nombre maximal d'itérations). Malheureusement, ce type de critère est purement fictif et peut donner de très

mauvais résultats en fonction de la population initiale et de la vitesse ou l'algorithme génétique va explorer l'espace.

Voici les critères les plus simples et les plus robustes expliqués dans la littérature :

- Le nombre de générations sans amélioration de la fonction objectif,
- Un critère original, basé sur la courbe de décroissance de la fonction objectif d'un algorithme utilisant un élitisme simple. Dans ce cas la décroissance est de type exponentielle et est approchée par une fonction. Les coefficients réels  $a$ ,  $b$  et  $c$  sont déterminés par régression non linéaire au sens des moindres carrés ou toute autre méthode d'interpolation exponentielle [MISS-1999]. La figure suivante présente la décroissance de la fonction objectif en fonction du nombre de génération. Elle montre aussi l'apport, sur la décroissance et sur l'optimum atteint, de la convergence locale (taux SQP correspond à la probabilité d'avoir une mutation de type déterministe d'une génération à une autre). Sans convergence locale, le minimum inférieur à 6 (p.u.) n'est pas atteint. Cette optimisation correspond à la minimisation des pertes dans un départ,



**Figure III- 6 : Evolution de la courbe d'approximation de la décroissance de la fonction objectif en fonction du type de mutation sélectionné**





# Chapitre IV :

Application et programmation des  
Algorithmes Génétiques pour résoudre  
notre problème

## IV-1 Enoncé du problème :

Notre souci est d'optimiser les coefficients d'un modèle de calcul d'une propriété donnée, ceci par l'élaboration d'un programme qui optimise les coefficients dans un domaine de recherche. Ce domaine varie d'une propriété à une autre, et dépend du modèle choisi.

### IV-1-1 Choix du modèle de chaque propriété :

Concernant notre étude, nous avons choisi des modèles pour chaque propriété étudiée comme suit :

- 1- pour la température critique :

$$T_{pc} = a \times T_{mav} + b \times 10^3 \times d_{20} + c$$

- 2- pour la pression critique :

$$P_{pc} = a \times T_{mav} + b \times 10^3 \times d_{20} + c$$

- 3- pour le volume critique :

$$V_{pc} = a \times T_{mav}^2 + b \times T_{mav} \times d_{20} + c$$

- 4- pour le facteur acentrique :

$$\omega = a \times T_{mav} \times d_{20} + c$$

Il est important de faire noter que nous avons essayé tout d'abord, d'appliquer le modèle développé pour les corps purs [17] aux fractions pétrolières qui a donné de bons résultats pour les températures et pressions critiques. Pour le volume critique et le facteur acentrique le modèle des corps purs a donné de mauvais résultats. Nous proposons dans cette étude deux nouveaux modèles.

### IV-1-2 Détermination de l'espace de recherche des coefficients de chaque modèle :

P \ coeff	a	b	d
Tc	[0,1]	[0,1]	[50,100]
Pc	[-1,1]	[-1,1]	[0,100]
Vc	[0,1]	[0,1]	[200,500]
W	[0,0.5]	[0,1]	--

Ce choix a été guidé par une estimation des valeurs des coefficients ainsi que par plusieurs tâtonnements des autres espaces de recherches possibles, cette partie de notre travail a pris beaucoup de temps avant de trouver l'espace de recherche des coefficients qui et des critères de la convergence du programme lors de son exécution.

## IV-2 organigramme de la programmation des A.G :

### IV-2-1 Codage réel des paramètres :

Pour résoudre notre problème nous avons choisi de coder les valeurs des coefficients de notre modèle (de chaque fraction pour chaque propriété) en codage réel, c'est-à-dire que les valeurs des éléments de notre matrice prendront des valeurs réelles qui appartiennent à l'espace de recherche définie. Parce que :

À l'aide du codage binaire, toutes les opérations sont assez simples à mettre en place. Malgré tout, quelques inconvénients existent :

– Il peut être difficile d'adapter ce codage à certains problèmes (notre problème):

La représentation binaire traditionnelle utilisée pour les algorithmes génétiques crée des problèmes pour les problèmes d'optimisation de grandes dimensions à haute précision numérique. Par exemple, avec 100 variables appartenant au domaine  $[-500; 500]$  et dont une précision de 6 chiffres après la virgule est requise, la taille du chromosome est 3000. Cela, en retour, génère un espace de recherche d'environ  $10^{1000}$ . Pour de tels problèmes, les algorithmes génétiques basés sur des représentations binaires ont de faibles performances.

– La distance de Hamming entre deux nombres réels proches peut être grand (exemple : 0111 qui vaut 7 et 1000 qui vaut 8, la distance est de 4). Ce qui crée bien souvent une convergence mais non l'obtention de la valeur optimale.

(La distance de Hamming entre deux chaînes de bits est le nombre de bits qui diffèrent de l'une à l'autre. Ainsi entre 01100101 et 00101100 la distance de Hamming vaut 3.)

– Suivant le problème, la résolution de l'algorithme peut être coûteux en temps.

– Le croisement et la mutation peuvent être inadaptés (création d'individus n'appartenant pas à l'espace de recherche).

Une des améliorations majeures consiste alors à se servir de nombres réels directement. Les résultats donnés par Michalewicz (1992) et Logan et Swaminathan (1994) montrent que la représentation binaire aboutit souvent à une moins bonne précision et qu'en règle général le gain en termes de temps de calcul. La conclusion à laquelle il arrive est naturellement qu'une représentation plus naturelle du problème offre des solutions plus efficaces.

En utilisant le codage réel, notre individu n'est alors plus qu'un chiffre à valeurs réelles dans l'espace des valeurs permises, L'opérateur de sélection reste identique à celui de la roue de loterie biaisée ou du tournoi.

### IV-2-2 Création de la population initiale :

Dans cette étape nous choisissons en premier lieu la population, cela varie surtout avec le but que l'on veut atteindre (l'erreur choisie). Cette valeur est prise comme étant le nombre de ligne d'une matrice que l'on appelle *pop*, le nombre de colonne de cette matrice est « m » (qui représente le nombre de coefficient). Ainsi nous allons retrouver avec une matrice de taille  $n*m$  tel que  $n$  : est la taille de population et  $m$  c'est le nombre de coefficient du modèle choisi, pour l'instant la matrice est vide, il faut générer des valeurs qui appartiennent à l'espace de recherche définie. Pour cela on fait appel à partir de notre programme principal à un sous programme (*genererpop*) qui va générer ces valeurs, par exemple si nous voulons générer une population de 04 coefficients ( $0 < a < 1$  et  $-5 < b < 5$  et  $0 < c < 50$  et  $0 < d < 100$ ) et de taille égale  $x$ , l'algorithme est suivant :

**Pour** 1 à  $x$

$$pop(i,1) = rand.$$

$$pop(i,2) = 10 * rand - 5.$$

$$pop(i,3) = 50 * rand.$$

$$pop(i,4) = 100 * rand.$$

$x$  : est la taille de la population.

Rand : une valeur tirée au hasard, elle est définie dans la banque de données du logiciel Matab. Comprise entre 0 et 1.

Donc le programme va générer des éléments de la matrice *pop* aléatoirement mais définie dans notre espace de recherche. Tel que la première colonne de la matrice *pop* représente le coefficient  $a$ , et la deuxième le coefficient  $b$ , et la troisième le coefficient  $c$  et en fin la quatrième colonne représente le coefficient  $d$ . Nous allons retrouver avec  $x$  corrélations ( $x$  étant la taille de la population) pour chaque propriété. Cette matrice générée est appelée population initiale. Par exemple  $x = 8$

Nous obtiendrons la matrice suivante :  $pop =$

$$\begin{pmatrix} 0.3412 & 0.3408 & 36.3557 & 30.9290 \\ 0.8385 & 0.6807 & 18.5207 & 70.2740 \\ 0.5466 & -0.5512 & 34.7284 & 62.1310 \\ 0.7948 & 4.5684 & 26.1295 & 88.0142 \\ 0.1730 & 4.7975 & 13.5724 & 25.2329 \\ 0.8757 & 2.3731 & 6.8259 & 01.1757 \\ 0.8939 & -3.0086 & 14.9362 & 66.1443 \\ 0.2844 & -0.3078 & 03.2391 & 98.8335 \end{pmatrix}$$

### IV-2-3 Calcul du fitness :

Pour le calcul de la fonction fitness (l'erreur de chaque ligne, c'est-à-dire de chaque corrélation), le programme exécute un sous programme qui est appelé à partir du programme principal, *fitness*, le rôle de ce programme est de calculer pour chaque ligne la propriété à partir du modèle et ceci pour chaque valeur de variable du modèle de la banque de donnée, et de la comparer avec celle du modèle.

**E<sub>i</sub> = 100\*valeur absolue (valeur du modèle – valeur de la banque de donnée)/ valeur de la banque de donnée.**

Pour une ligne donnée, notre fonction calcule autant de fois E que la taille du vecteur de notre banque de donnée, et la valeur du fitness de cette ligne sera égale à Max(E).

Nous aurons alors, pour chaque ligne, une erreur maximale pour la corrélation proposée. Par exemple pour une ligne quelconque, si nous avons une banque de donnée de « n » valeurs de Y correspondante à « n » valeurs de x et z nous prendrons la valeur du fitness qui correspond au maximum des valeurs E, donc la valeur du fitness de cette ligne sera égale à :

$$fitness(i) = \max(E_1 E_2 E_3 \dots E_n)$$

**I** : représente la ligne de la matrice.

**Max** : une fonction définie dans le Matlab qui donne la valeur maximale.

Ainsi pour chaque corrélation (ligne) nous aurons une erreur maximale. Enfin les valeurs de cet fitness seront introduites dans un vecteur.

### IV-2-4 Opérateur de sélection :

C'est le premier opérateur exécuté par le programme principal, ainsi il fait appel à un sous programme (*sélection*) qui sélectionnant les meilleurs individus pour les deux autres opérateurs qui vont suivre. Dans ce sous programme on calcule en premier lieu deux probabilités « p » et « q » qui définit comme suit :

$$p(i) = \frac{fitness(i)}{\sum_{j=1}^x fitness(j)} \quad q(i) = \sum_{j=1}^i p(j).$$

**i** : l'indice de la ligne

**x** : la taille de la population

Le résultat de ces deux probabilités est reporté sur deux vecteurs que l'on appelle  $p_s$  et  $q_s$ .

Après cette étape nous générons un vecteur  $r_s$  de taille égale à la taille de la population choisie, ce vecteur est constitué d'éléments rand. Chaque élément de ce vecteur va être borné par les probabilités  $q$  que nous avons reportées sur le vecteur  $q_s$ , le programme choisit

l'indice de la borne supérieur et copie la ligne de la *pop* qui correspond à cet indice (c'est la ligne choisi), sur la ligne qui correspond à l'indice de l'élément du vecteur  $r_s$  borné (c'est la ligne qui va disparaître), cette opération s'exécute seulement si le fitness de la ligne choisi est inférieur à celle qui va disparaître, sinon il va choisi la borne supérieur suivante.

**L'algorithme de cette fonction est :**

**Pour i = 1 à x**

**Pour j = 1 à x**

**Si  $r_s(i) \leq q(j)$  alors**

**Si  $fitness(j) < fitness(i)$  alors**

**Pour k = 1 à 3**

$pop(i, k) = pop(j, k)$

**Fin**

**Sinon choisir la borne suivante**

**Fin**

**Fin**

**Fin**

**Fin**

nous obtiendrons une nouvelle matrice *pop'*

#### **IV-2-5 Opérateur de croisement :**

Cet opérateur vise à améliorer le fitness, c'est-à-dire l'erreur de chaque individu, en combinant deux coefficients de la même colonne, de deux individus différents, pour ce faire notre programme principal fait appel à un sous programme (*croisement*) qui va croiser soit en un point ou en deux points ou plus, tout dépendra de la probabilité  $p_c$ , ou du nombre de coefficients que nous avons, dans notre cas, l'exécution se fait au maximum en deux points.

Nous définissons tous d'abord la probabilité de croisement  $p_c$ , c'est une probabilité qui est fixée généralement à 0.75 et qui va choisir les individus qui vont subir un croisement. Dans notre cas nous n'avons pas eu recours à la modifier mais il est possible que pour d'autres types de modèles, nous serons obligés de la modifier soit en l'augmentant ou en la diminuant.

Le sous programme choisit tous d'abord les individus qui vont subir un croisement, et cela en créant un vecteur que l'on appelle  $r_c$ , ce vecteur est composé d'éléments rand c'est-à-dire pris aléatoirement entre [0,1] tel que le nombre d'éléments soit égal à la taille de la population. Ensuite il va choisir les individus qui sont aptes pour un éventuel croisement,

c'est-à-dire il va tester chaque élément de  $r_c$  et si ce dernier est inférieur à 0.75 alors il prend l'indice de cet élément et cet indice représente la ligne de l'individu ainsi choisi. Après la sélection des individus pour une éventuelle opération de croisement, notre sous programme vérifie le nombre des individus choisis, si le nombre est impair il laisse tomber le dernier individu choisi puisque notre croisement s'effectue deux par deux (père, mère). Ensuite il va prendre deux individus successifs et leurs faire subir un croisement. Mais avant il va sélectionner les colonnes qui vont être croisées, ce choix est lui aussi régi par une probabilité  $p_c$ . Pour chaque couple d'individus, un autre vecteur  $r'_c$  est créé composé de quatre éléments rand (quatre est le nombre de nos coefficients). Chaque élément est testé ; si le premier élément est inférieur à  $p_c$  alors la première colonne des deux individus (père, mère) est permutée, et ainsi de suite pour les deux autres colonnes.

#### IV-2-6 Opérateur de Mutation :

C'est le dernier opérateur que le programme principal exécute, et c'est probablement le plus important puisque c'est lui qui fait évoluer la génération. La mutation s'exécute par un sous programme que l'on appelle *mutation*, et qui choisit un des éléments de la matrice  $pop^n$  à pour être muté, dans le domaine de recherche fixé auparavant.

Dans ce sous programme, nous choisissons tous d'abord l'élément qui va être muté, ce choix est guidé par une probabilité de mutation  $p_m$ , cette probabilité varie d'un problème à un autre. Pour notre exécution nous l'avons fixé à 0.33. Donc le programme va créer un vecteur  $r_m$  d'élément rand tel que la taille de ce vecteur soit égale à : taille de la population \* nombre du coefficient. Ce vecteur est important, vu que c'est lui qui va choisir les éléments qui vont être mutés, et cela en testant chaque élément de ce vecteur, si l'élément est inférieur à notre probabilité de mutation  $p_m$ , alors l'élément de notre matrice  $pop^n$  qui correspond à l'indice de l'élément de ce vecteur est choisi, sinon il passe à l'élément suivant du vecteur. Pour trouver l'élément de la matrice  $pop^n$  qui correspond à l'indice de l'élément de ce vecteur, le programme exécute un calcul de ligne et de la colonne d'après cette formule.

$$j = \text{mod}(k, m)$$

$$i = 1 + ((k - j) \div m)$$

J : l'indice de la colonne de la matrice  $pop^n$  (ne peut être > m).

i : l'indice de la matrice de la matrice  $pop^n$

k : l'indice de l'élément du vecteur  $r_m$

$m$  : le nombre de la colonne de la matrice  $pop^n$  (nombre du coefficient).

$\text{mod}$  : c'est une fonction qui existe dans la banque de donnée de matlab et donne le reste de la division de  $k/m$

Après la sélection des éléments à muter, nous passons à l'étape de la mutation proprement dite. Cette opération est très simple vue que nous connaissons maintenant la colonne et la ligne de l'élément à muter. La mutation de l'élément est une simple recherche d'une éventuelle valeur dans son intervalle. Ainsi pour tous les éléments choisis, on effectue la même opération.

**L'algorithme du sous programme de la mutation pour ce modèle :**

**Pour**  $k = 1$  à  $x*m$

**Si**  $r_m < 0.33$  **alors**

$j = \text{mod}(k, m)$

**Si**  $j = 0$  **alors**

$j = m$

**Fin**

$i = 1 + ((k - j) \div m)$ .

**Si**  $j > 1$  **alors**

**Si**  $j > 2$  **alors**

$pop^n(i, j) = 100 * rand$

**Sinon**

$pop^n(i, j) = 50 * rand$

**Fin**

**Sinon**

$pop^n(i, j) = 10 * rand - 5$

**Fin**

**Sinon**

$pop^n(i, j) = rand$

**Fin**

**Fin**

**Fin**



### IV-2-7 Test d'arrêt :

Après l'exécution de toutes ces opérations nous obtenons une nouvelle génération d'individus. Le test d'arrêt est une simple vérification du fitness de chaque ligne de notre matrice, notre programme repère le minimum du vecteur fitness et si ce dernier est inférieur à l'erreur choisi au début (que l'on veut atteindre) le programme s'arrête et nous donne la ligne, et exécute ainsi un autre sous programme qui fait seulement appliquer le modèle ainsi obtenu. Sinon le programme refait tous les calculs avec les différents opérateurs mais cette fois à avec une matrice initiale *pop*". Afin d'améliorer, nos résultats nous pouvons poser deux tests d'arrêts c'est-à-dire deux conditions pour que le programme s'arrête (l'erreur maximale et l'erreur moyenne).

### IV-3 Exemple mettant en œuvre tous les opérations :

Pour comprendre l'enchaînement des outils définis ci-dessus, la meilleure solution est probablement un exemple. Supposons, bien que cet exemple soit simpliste, que l'on veuille optimiser les coefficients de la fonction  $f(x, y) = a*x^2 + b*x*y + c$ .

Avec une banque de données de trois vecteurs (matrice de 1 ligne et 49 colonne)

Un vecteur pour les valeurs de la fonction  $f(x, y)$ .

Un vecteur pour les valeurs de  $x$ .

Un vecteur pour les valeurs de  $y$ .

Cet exemple suivant va permettre de mettre en oeuvre sur une génération tous les opérateurs étudiés de manière à pouvoir observer le résultat de l'application successive de ces opérateurs. Il consiste à trouver les coefficients de la fonction objectifs.

On initialise la population initiale de manière aléatoire et on fixe sa taille à 6 individus. On définit simplement la fitness comme étant la valeur de l'erreur, vu qu'on en cherche la valeur minimum, plus la valeur de *fitness* sera minime plus on se rapprochera des coefficients de la fonction identité.

Nous aurons donc à générer une population de «  $n$  » individus à trois coefficients, avec un espace de recherche définie comme suit pour les trois coefficients ;  $0 < a < 1$ ,  $0 < b < 1$ ,  $20 < c < 25$ . Bien sur le coefficient (**a**) représente la colonne numéro un dans notre matrice pop et (**b**) la deuxième colonne, et (**c**) la troisième colonne. Nous suivons maintenant les différentes étapes de l'algorithme

a- Tirage de la population initiale que nous limitons à 6 individus pour faciliter la compréhension. Après ce tirage, la fonction est évaluée pour chacun des individus

	chromosome	fitness
Individu 1	0.015 0.75 21.33	112.07
Individu 2	0.93 0.46 21.25	41.388
Individu 3	0.84 0.52 20.60	57.06
Individu 4	0.68 0.30 21.62	6.143
Individu 5	0.67 0.84 20.05	142.79
Individu 6	0.15 0.69 21.13	99.535

b- population sélectionnée par la reproduction. À partir de la population initiale et de la simulation de la roue de loterie, la population issue de la reproduction peut être la suivante les individus les plus « forts » y sont bien sûr présents

	chromosome	fitness
Individu 1	<b>0.68 0.30 21.62</b>	<b>6.143</b>
Individu 2	0.93 0.46 21.25	41.38
Individu 3	<b>0.68 0.30 21.62</b>	<b>6.143</b>
Individu 4	<b>0.68 0.30 21.62</b>	<b>6.143</b>
Individu 5	0.67 0.84 20.05	142.79
Individu 6	<b>0.68 0.30 21.62</b>	<b>6.143</b>

c- population après l'opération de croisement (la probabilité de croisement 0.75) les individus 1 et 2 forment le premier couple de l'opération de croisement, mais suite à un tirage aléatoire ils ne se croisent pas, les individus 3 et 4 et 5 et 6 se croisent avec comme site de croisement le 2 (ce nombre est lui aussi issu d'un tirage aléatoire).

La population après croisement est alors la suivante :

	chromosome	fitness
Individu 1	0.68 0.30 21.62	6.143
Individu 2	0.93 0.46 21.25	41.38
Individu 3	0.68 0.30 21.62	6.143
Individu 4	0.68 0.30 21.62	6.143

Individu 5	0.67	<b>0.30</b>	20.05	5.87
Individu 6	0.68	<b>0.84</b>	21.62	145.42

**d-** population après l'opération de mutation.

L'opération de mutation intervient suivant un tirage aléatoire, biaisé par une probabilité fixée par l'utilisateur ici pour montre son influence, nous supposons qu'un unique bit mute le premier du 5<sup>ème</sup> individu.

	chromosome			fitness
Individu 1	0.68	0.30	21.62	6.143
Individu 2	0.93	0.46	21.25	41.38
Individu 3	0.68	0.30	21.62	6.143
Individu 4	0.68	0.30	21.62	6.143
Individu 5	<b>0.52</b>	0.30	20.05	<b>1.23</b>
Individu 6	0.68	0.84	21.62	145.42

Population de la seconde génération. A la suite de l'application des trois opérateurs ci-dessus, la nouvelle population est la suivante :

	chromosome			fitness
Individu 1	0.93	0.30	21.62	36.49
Individu 2	0.68	0.46	21.25	12.13
Individu 3	0.52	0.30	20.05	1.23
Individu 4	0.68	0.30	21.62	6.143
Individu 5	0.52	0.30	20.05	1.23
Individu 6	0.52	0.30	20.05	1.23

On voit sur cet exemple que la performance bien, mais la meilleure ont augmenté lors du changement de génération, si l'algorithme génétique réalise plusieurs itérations de cette procédure, les solution tendront vers un résultat optimal, ce qui peut être démontrés par le théorème fondamental sur les schémas. néanmoins, ici la petite taille de la population fait que de nombreuses itérations peuvent être nécessaires à l'obtention d'un résultat optimal, ce qui ne serait pas le cas si la population était plus vaste

# Chapitre V

## Résultats et commentaires

## V.1 Introduction :

L'exécution du programme principal s'effectue dans le command window du matlab, pour l'exécution de notre programme, on fait appel à notre programme principal en tapant *pro-critique*, lors de l'exécution nous devons choisir la propriété ( $T_{pc}$  ou  $P_{pc}$  ou  $V_{pc}$  ou  $\omega$ ).

L'exécution vise à minimiser l'erreur maximale pour toutes les fractions en même temps c'est-à-dire une corrélation pour les **49** fractions. Pour atteindre notre but il nous a fallu une population initiale de **3000** individus.

## V.2 Résultats obtenus :

### V.2-1 Température critique :

L'erreur maximum qu'on peut enregistrer pour la température critique une erreur maximale de **1.5 %** et ceci après un grand nombre d'essais d'exécution du programme : le programme converge après **16** générations et un temps d'exécution inférieur à **2** minutes, l'erreur maximale obtenue est de **1.2021**, l'erreur moyenne est de **0.5364** et enfin l'erreur minimale est **0.0086**

Le modèle obtenu par les algorithmes génétiques est suivant :

$$T_{pc} (^{\circ}\text{k}) = 0.6209 \times T_{\text{nav}} (^{\circ}\text{C}) + 551.3 \times d_{20} + 97.7829.$$

L'application du programme proposé consiste à donner la température d'ébullition et la densité de la fraction choisie pour obtenir directement sa température critique.

Les résultats obtenus et les valeurs calculées à partir des autres corrélations sont représentés sur le tableau **V-1**

$T_{pc}$	$T_{pc1}$	$e_1$	$T_{pc2}$	$e_2$	$T_{pc3}$	$e_3$	$T_{pc4}$	$e_4$
502.56	496.7069	1.1647	531.89	5.8361	501.01	0.3084	555.72	10.5778
512.89	511.6644	0.2390	535.23	4.3557	509.95	0.5732	562.76	9.7233
518.85	517.9113	0.1809	540.41	4.1553	515.85	0.5782	569.72	9.8044
525.66	525.8672	0.0394	545.15	3.7077	522.38	0.6240	576.60	9.6907
533.15	535.2014	0.3848	549.55	3.0761	529.45	0.6940	583.40	9.4251
539.50	542.2753	0.5144	554.53	2.7859	535.69	0.7062	590.13	9.3846
545.64	548.9081	0.5989	559.44	2.5291	541.75	0.7129	596.77	9.3706
551.29	554.6037	0.6011	564.50	2.3962	547.42	0.7020	603.34	9.4415
556.94	560.2993	0.6032	569.54	2.2624	553.08	0.6931	609.84	9.4983
562.91	566.6565	0.6656	574.41	2.0430	559.01	0.6928	616.26	9.4775
568.76	572.7931	0.7091	579.25	1.8444	564.84	0.6892	622.61	9.4680
574.44	578.5990	0.7240	584.10	1.6816	570.51	0.6841	628.88	9.4771

580.19	584.5703	0.7550	588.91	1.5030	576.24	0.6808	635.08	9.4607
585.72	590.1005	0.7479	593.74	1.3693	581.77	0.6744	641.22	9.4755
591.31	595.7961	0.7587	598.5	1.2159	587.36	0.6680	647.28	9.4654
596.02	599.7827	0.6313	603.6	1.2718	592.15	0.6493	653.27	9.6054
602.77	607.7386	0.8243	607.81	0.8361	598.77	0.6636	659.20	9.3618
608.15	613.0483	0.8054	612.54	0.7219	604.15	0.6577	665.06	9.3579
613.37	618.0824	0.7683	617.27	0.6358	609.38	0.6505	670.86	9.3728
618.10	622.1792	0.6600	622.16	0.6569	614.14	0.6407	676.59	9.4629
621.82	624.4016	0.4152	627.50	0.9134	617.94	0.6240	682.25	9.7182
627.10	629.6011	0.3988	632.09	0.7957	623.18	0.6251	687.85	9.6875
632.14	634.3595	0.3511	636.73	0.7261	628.19	0.6249	693.39	9.6893
637.19	639.1730	0.3112	641.31	0.6466	633.21	0.6246	698.87	9.6800
642.47	644.4276	0.3047	645.75	0.5105	638.44	0.6273	704.29	9.6222
647.60	649.4065	0.2789	650.21	0.4030	643.52	0.6300	709.65	9.5815
652.33	653.6687	0.2052	654.80	0.3786	648.21	0.6316	714.96	9.6010
656.87	657.6002	0.1112	659.42	0.3882	652.70	0.6348	720.20	9.6412
661.67	662.0278	0.0541	650.05	1.7562	657.44	0.6393	725.39	9.6302
666.61	666.7310	0.0182	654.35	1.8392	662.31	0.6451	730.52	9.5873
671.36	674.2080	0.4242	658.73	1.8813	666.98	0.6524	735.60	9.5686
679.25	687.4185	1.2026	661.41	2.6264	674.81	0.6537	740.63	9.0364
687.38	693.7378	0.9249	670.68	2.4295	682.77	0.6707	750.52	9.1856
696.90	702.6482	0.8248	679.12	2.5513	692.11	0.6873	760.21	9.0845
706.35	711.5034	0.7296	687.45	2.6757	701.37	0.7050	769.71	8.9701
716.48	721.6267	0.7183	695.24	2.9645	711.32	0.7202	779.01	8.7274
724.58	728.2216	0.5026	704.10	2.8265	719.15	0.7494	788.14	8.7720
733.83	736.9114	0.4199	712.14	2.9557	728.16	0.7727	797.10	8.6219
742.25	744.2230	0.2658	720.57	2.9208	736.29	0.8030	805.90	8.5753
750.53	751.3692	0.1118	728.99	2.8700	744.25	0.8367	814.52	8.5260
761.66	763.4220	0.2313	735.39	3.4490	755.24	0.8429	823.01	8.0548
769.77	770.4028	0.0822	743.71	3.3854	763.01	0.8782	831.35	7.9998
778.64	778.7068	0.0086	751.39	3.4997	771.59	0.9054	839.55	7.8226
787.70	787.3966	0.0385	758.80	3.6689	780.37	0.9306	847.63	7.6082
796.77	796.1416	0.0789	766.09	3.8505	789.17	0.9539	855.59	7.3823
804.97	800.4038	0.5672	773.84	3.8672	797.03	0.9864	863.44	7.2636

812.87	804.2250	1.0635	781.73	3.8309	804.54	1.0248	871.18	7.1733
817.27	808.4321	1.0814	785.33	3.9081	808.79	1.0376	875.02	7.0662
821.72	812.7495	1.0917	788.83	4.0026	813.10	1.0490	878.83	6.9501

$T_{pc}$  : La base de donnée (corrélacion de Riazi-Daubert)

$T_{pc1}$  : Corrélacion proposée.  $e_i$  : Erreur par rapport à la corrélacion Riazi-Daubert

$T_{pc2}$  : Corrélacion de Skander.

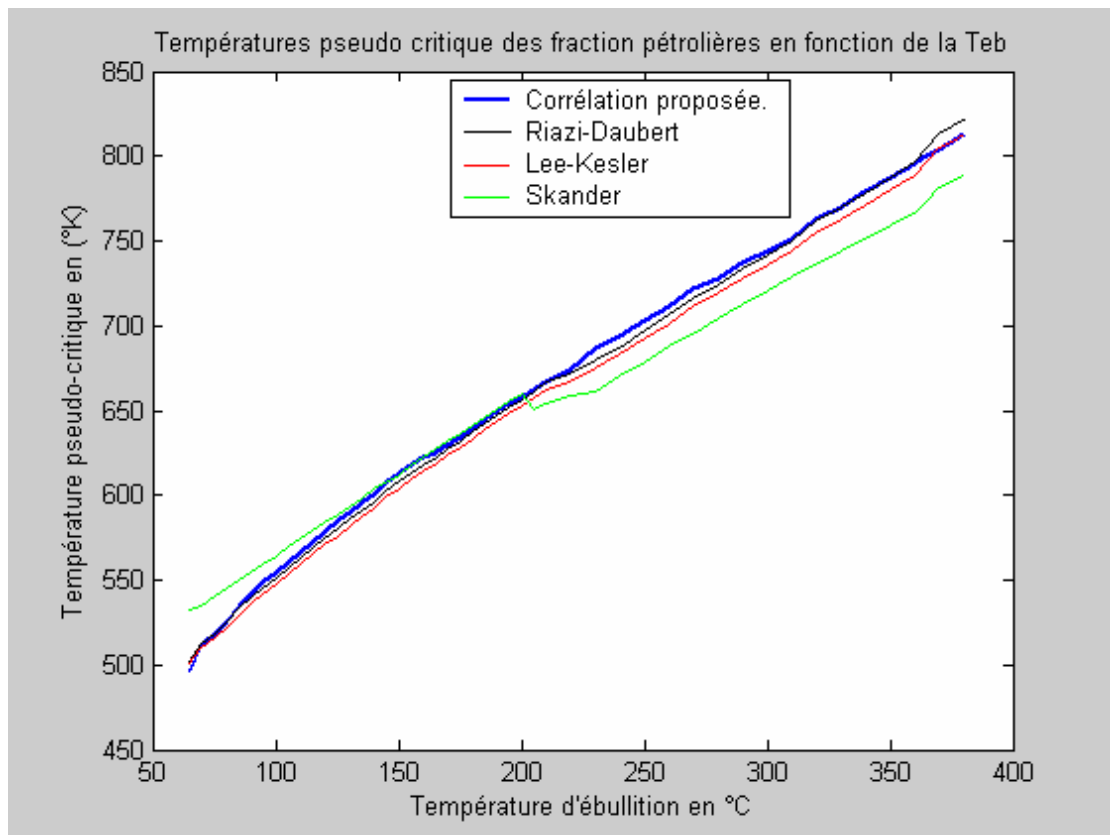
$T_{pc3}$  : Corrélacion de Lee-Kesler.

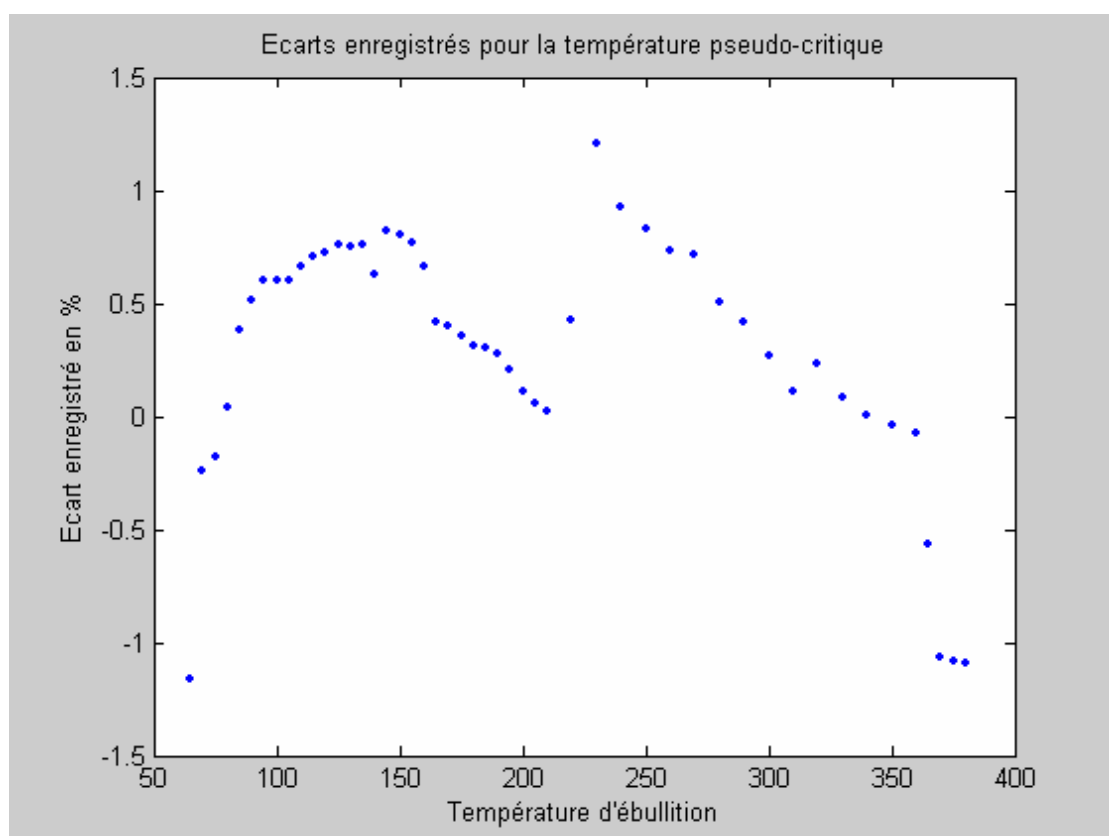
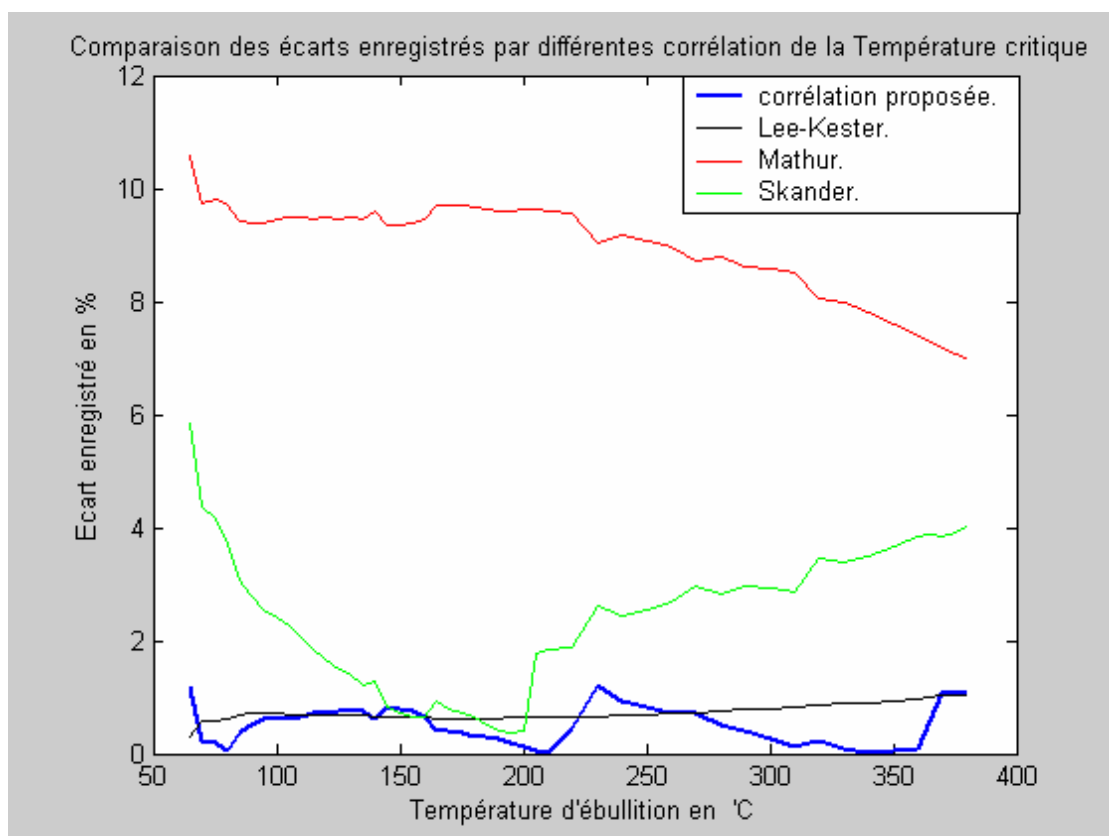
$T_{pc4}$  : Corrélacion de Mathur.

**V.2-1-2 Etude graphique :**

Nous avons représenté dans les figures suivantes les valeurs de la banque de donnée et celle trouvées par le modèle proposé par les algorithmes génétiques et les autres corrélations.

Nous avons aussi représenté l'erreur pour chaque corrélacion de chaque valeur de  $T_{pc}$







### V.2-2-1 Pression critique :

L'erreur maximum qu'on peut enregistrer pour la pression critique une erreur maximale de **5 %** et ceci après un grand nombre d'essais d'exécution du programme : le programme converge après **39** générations et un temps d'exécution inférieur à **6** minutes, l'erreur maximale obtenue est de **4.9789**, l'erreur moyenne est de **2.5428** et enfin l'erreur minimale est **0.0891**

Le modèle obtenu par les algorithmes génétiques est suivant :

$$P_{pc} (bar) = -0.045 \times T_{mav} (°C) - 18.6 \times d_{20} + 45.4047$$

L'application du programme proposé consiste à donner la température d'ébullition et la densité de la fraction choisie pour obtenir directement sa pression critique.

Les résultats obtenus et les valeurs calculées à partir des autres corrélations sont représentés sur le tableau V-2

$P_{pc}$	$P_{pc1}$	$e_1$	$P_{pc2}$	$e_2$	$P_{pc3}$	$e_1$
28.9200	30.3823	5.0562	28.4000	1.7981	30.7200	1.1678
30.1500	29.7574	1.3023	29.8000	1.1609	31.3500	5.2824
29.7300	29.4263	1.0214	29.5000	0.7736	31.0100	5.3268
29.7300	29.0377	2.3288	29.5300	0.6727	30.8300	6.0287
29.8000	28.6025	4.0185	29.8000	0	30.7900	7.3407
29.5600	28.2436	4.4535	29.6400	0.2706	30.5400	7.7687
29.2500	27.8995	4.6170	29.4000	0.5128	30.2200	7.9333
28.7800	27.5871	4.1449	29.1100	1.1466	29.8100	7.7238
28.3400	27.2747	3.7591	28.7000	1.2703	29.4000	7.4994
28.0200	26.9399	3.8546	28.5000	1.7131	29.0700	7.6019
27.6700	26.6126	3.8213	28.2000	1.9154	28.7100	7.5799
27.2800	26.2965	3.6052	27.9000	2.2727	28.3100	7.3809
26.9300	25.9748	3.5470	27.6000	2.4879	27.9500	7.3346
26.5300	25.6679	3.2494	27.2000	2.5254	27.5400	7.0564
26.1600	25.3555	3.0752	26.9000	2.8287	27.1500	6.8596
25.5600	25.1008	1.7967	26.4000	3.2864	26.6300	5.9829
25.5400	24.7121	3.2417	26.4000	3.3673	26.4500	6.8047
25.1500	24.4127	2.9317	26.1000	3.7773	26.0400	6.4705
24.7300	24.1226	2.4562	25.7000	3.9224	25.6200	6.0551
24.2100	23.8641	1.4287	25.2000	4.0892	25.1300	5.2288

23.4600	23.6689	0.8903	24.4000	4.0068	24.5100	3.5854
23.1200	23.3732	1.0951	24.1000	4.2388	24.1300	3.2734
22.7300	23.0924	1.5943	23.7000	4.2675	23.7200	2.7612
22.3600	22.8097	2.0113	23.3000	4.2039	23.3200	2.2821
22.0500	22.5122	2.0961	23.0000	4.3084	22.9600	2.0309
21.7200	22.2239	2.3202	22.7000	4.5120	22.5800	1.6393
21.3200	21.9599	3.0013	22.3000	4.5966	22.1600	0.9386
20.8900	21.7070	3.9109	21.9000	4.8348	21.7200	0.0623
20.5400	21.4373	4.3687	21.5000	4.6738	20.1300	6.3648
20.2200	21.1584	4.6409	21.2000	4.8467	19.7900	6.7676
19.8800	20.6656	3.9518	20.8000	4.6278	19.4300	6.2154
20.1400	19.9794	0.7974	21.2000	5.2632	19.4000	2.8769
19.2400	19.5257	1.4848	20.2000	4.9896	18.5900	4.8632
18.6400	18.9845	1.8484	19.5000	4.6137	17.9400	5.6038
18.0700	18.4453	2.0767	18.9000	4.5932	17.3200	6.2272
17.6400	17.8632	1.2653	18.5000	4.8753	16.7700	6.1973
16.9300	17.4002	2.7772	17.6000	3.9575	16.0900	7.7388
16.4400	16.8665	2.5942	17.1000	4.0146	15.5200	8.1903
15.8600	16.3793	3.2741	16.4000	3.4048	14.9100	9.2641
15.3000	15.8977	3.9063	15.7000	2.6144	14.3200	10.3115
15.1400	15.2505	0.7299	15.5000	2.3778	13.9300	8.7219
14.6100	14.7745	1.1257	14.9000	1.9849	13.3800	9.5446
14.2000	14.2538	0.3787	14.4000	1.4085	12.8900	9.6041
13.8400	13.7201	0.8665	14.0000	1.1561	12.4200	9.3936
13.5000	13.1845	2.3369	13.5000	0	11.9800	8.9224
13.0800	12.9205	1.2197	13.0000	0.6116	11.5000	10.8598
12.6600	12.6713	0.0891	12.5000	1.2638	11.0300	12.9643
12.5000	12.4091	0.7274	12.3000	1.6000	10.8200	12.7126
12.3600	12.1432	1.7544	12.1000	2.1036	10.6200	12.3233

$P_{pc}$  : La base de donnée (corrélacion de Riazi-Daubert).

$P_{pc1}$  : Corrélacion proposée.

$e_i$  : Erreur par rapport à la corrélacion de Riazi-Daubert

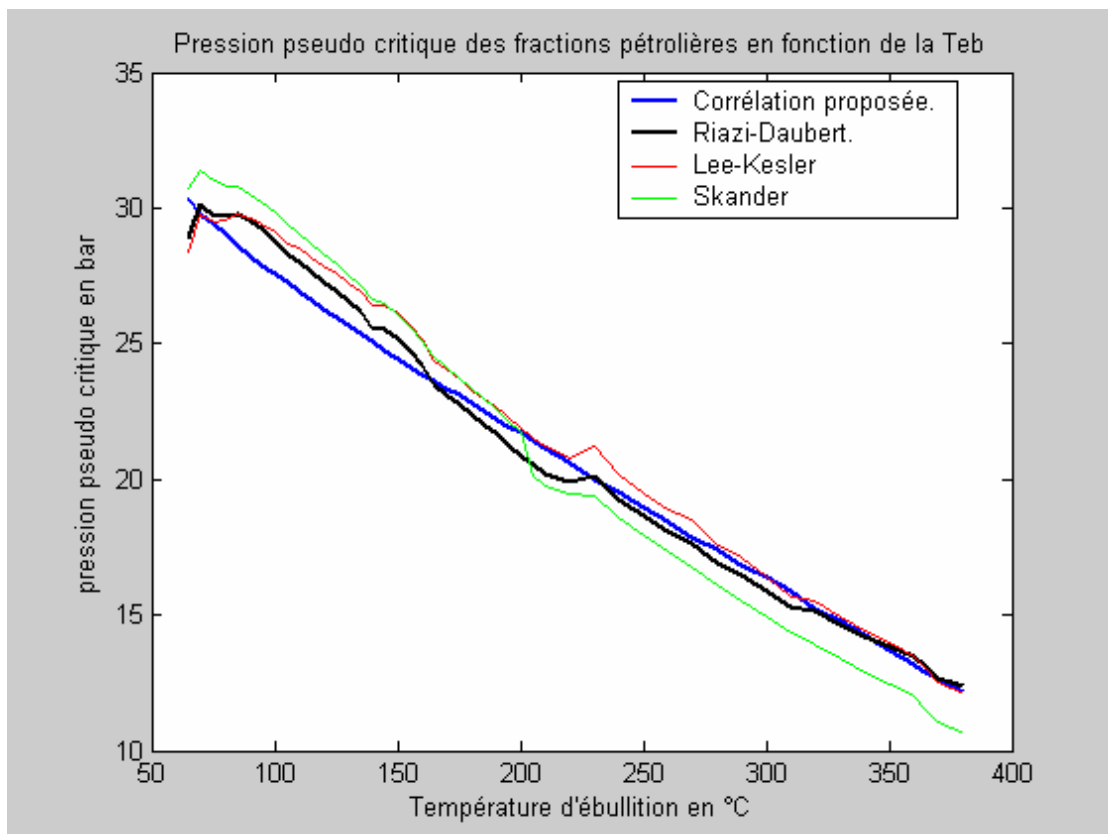
$P_{pc2}$  : Corrélacion de Lee-Kesler.

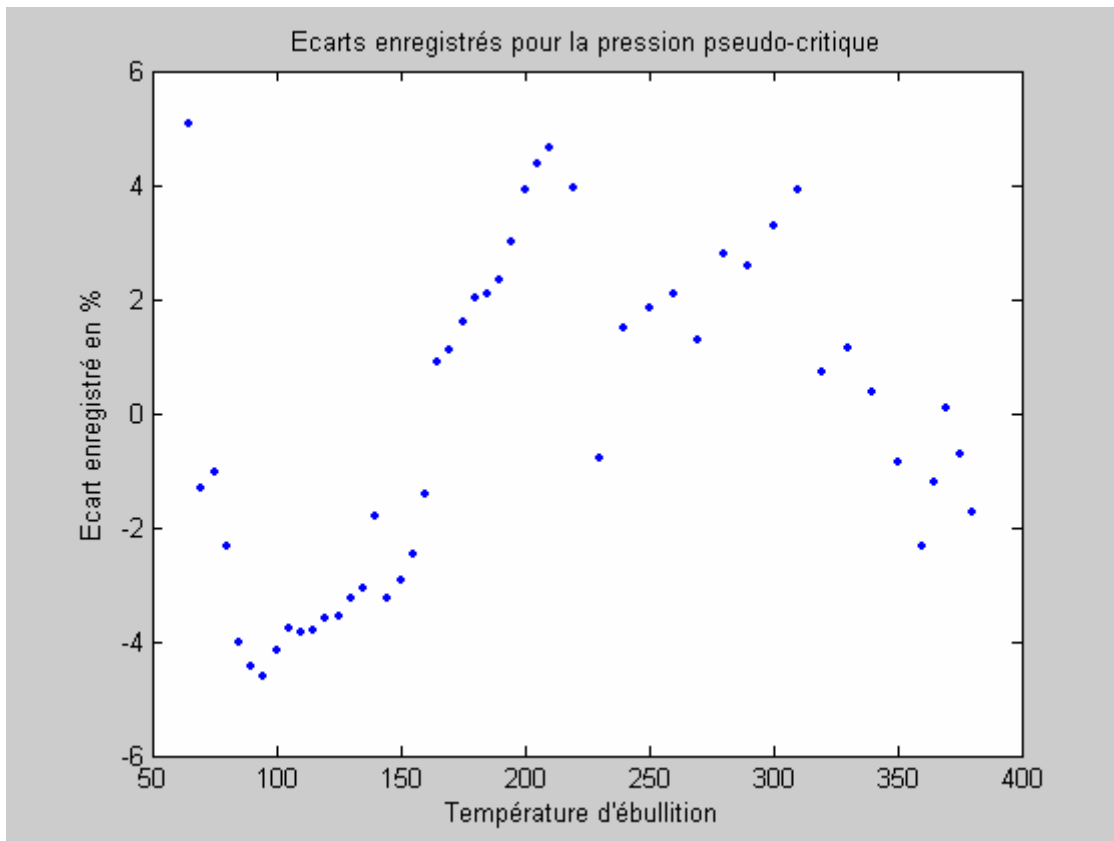
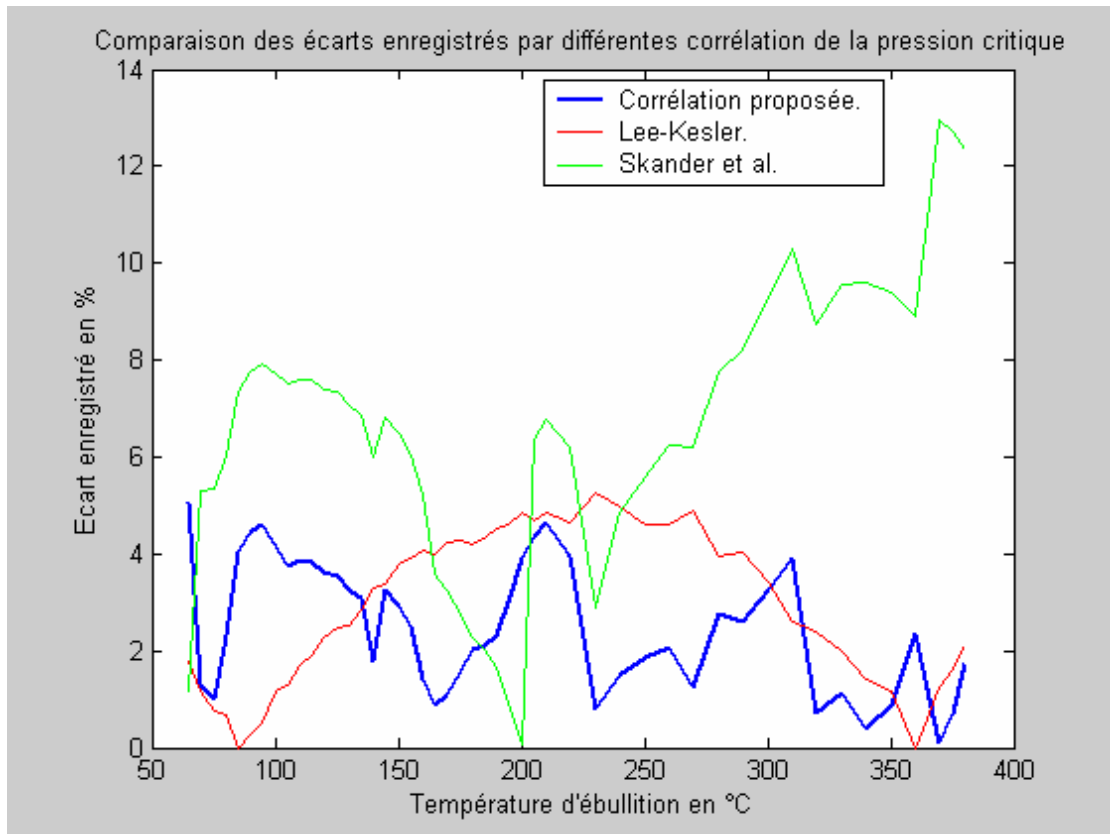
$P_{pc3}$  : Corrélacion de Skander et al.

### V.2-2-2 Etude graphique :

Nous avons représenté dans les figures suivantes les valeurs de la banque de donnée et celle trouvées par le modèle proposé par les algorithmes génétiques et les autres corrélations.

Nous avons aussi représenté l'erreur pour chaque corrélation de chaque valeur de  $P_{pc}$





**V-2-3-1 Volume critique :**

L'erreur maximum qu'on peut enregistrer pour le volume critique une erreur maximale de **3 %** et ceci après un grand nombre d'essais d'exécution du programme : le programme converge après **12** générations et un temps d'exécution inférieur à **2** minutes, l'erreur maximale obtenue est de **2.5523**, l'erreur moyenne est de **0.8063**, et enfin l'erreur minimale est **0.0099**.

Le modèle obtenu par les algorithmes génétiques est le suivant :

$$V_{pc} \text{ (cm}^3 \text{ / g.mol)} = 0.0038 \times T_{mav}^2 \text{ (}^\circ\text{C)} + 0.9814 \times T_{mav} \times d_{20} + 310.6783$$

L'application du programme proposé consiste à donner la température d'ébullition et la densité de la fraction choisie pour obtenir directement ce volume critique.

Les résultats obtenus et les valeurs calculées à partir des autres corrélations sont représentés sur le tableau **V-3**

$V_{pc}$	$V_{pc1}$	$e_1$	$V_{pc2}$	$e_2$	$V_{pc3}$	$e_1$
377.98	368.2	2.5523	383.05	3.8708	388.1	5.1216
383.90	375.5	2.1994	379.33	1.0211	400.55	6.2648
385.16	381.9	0.8391	386.90	1.2851	413.01	7.5257
388.94	388.9	0.0131	392.32	0.87456	425.47	8.5978
394.39	396.3	0.4931	396.01	0.082002	437.92	9.4961
400.97	403.7	0.6857	402.54	0.29296	450.38	10.36
408.35	411.3	0.7200	409.78	0.36847	462.84	11.138
416.31	418.9	0.6333	418.30	0.1545	475.29	11.855
424.72	426.8	0.4990	426.87	0.0071982	487.75	12.488
433.49	435.1	0.3709	434.61	0.11221	500.21	13.017
442.56	443.6	0.2257	442.73	0.18721	512.66	13.479
451.88	452.2	0.0693	451.37	0.1824	525.12	13.888
461.43	461.1	0.0709	459.82	0.27895	537.58	14.226
471.18	470.2	0.2184	468.98	0.2497	550.03	14.523
481.12	479.5	0.3425	477.96	0.3164	562.49	14.759
491.25	488.6	0.5387	489.50	0.1831	574.95	15.018
501.56	499.0	0.5173	495.30	0.74006	587.40	15.055
512.04	508.9	0.6138	505.01	0.76969	599.86	15.164
522.68	519.0	0.7076	515.28	0.71839	612.32	15.243

533.49	529.0	0.8372	527.04	0.3764	624.77	15.325
544.47	538.7	1.0555	541.80	0.56792	637.23	15.459
555.61	549.5	1.1002	552.10	0.4713	649.69	15.422
566.90	560.4	1.1533	563.20	0.50395	662.14	15.371
578.36	571.5	1.1925	574.32	0.49742	674.60	15.289
589.98	582.9	1.1949	584.79	0.31801	687.06	15.156
601.75	594.5	1.1994	595.81	0.21442	699.51	15.007
613.68	606.1	1.2337	608.12	0.33066	711.97	14.869
617.84	617.8	0.0099	621.14	0.54114	724.43	14.722
630.79	629.8	0.1515	653.48	3.6185	736.88	14.527
643.47	642.2	0.1964	665.70	3.5292	749.34	14.297
669.27	666.8	0.3616	678.58	1.7287	761.80	12.464
695.64	694.6	0.1429	681.47	1.9335	774.26	10.283
722.58	720.5	0.2868	711.89	1.2105	786.71	8.4151
750.10	748.3	0.2419	738.14	1.3745	799.17	6.3671
778.18	776.9	0.1651	764.91	1.5668	811.63	4.2797
806.82	807.0	0.0183	789.66	2.1918	824.08	2.0765
836.03	836.2	0.0181	821.65	1.7685	836.54	0.0429
865.79	867.2	0.1686	850.05	2.0234	849.01	2.1484
896.11	898.4	0.2590	881.72	1.8952	861.45	4.2928
926.98	930.3	0.3603	914.27	1.7555	873.91	6.4549
958.40	965.8	0.7718	936.92	3.0821	886.37	8.9609
990.36	999.3	0.8989	970.59	2.9542	898.82	11.175
1022.88	1034.3	1.1181	1001.90	3.2356	911.28	13.502
1055.94	1070.4	1.3738	1032.9	3.635	923.74	15.882
1089.54	1175	1.6446	1064.1	4.0747	998.48	10.914
1106.54	1126.2	1.7779	1099.1	2.4668	1010.93	11.404
1123.68	1144.9	1.8872	1135.66	0.81243	1023.39	11.872
1140.95	1164.0	2.0220	1152.21	1.025	1035.85	12.373
1158.36	1183.4	2.1650	1168.79	1.2533	1048.3	12.891

$V_{pc}$  : La base de donnée (corrélacion de Riazi-Daubert).

$V_{pc1}$  : Corrélacion proposée.

$e_i$  : Erreur par rapport à la corrélacion de Riazi-Daubert

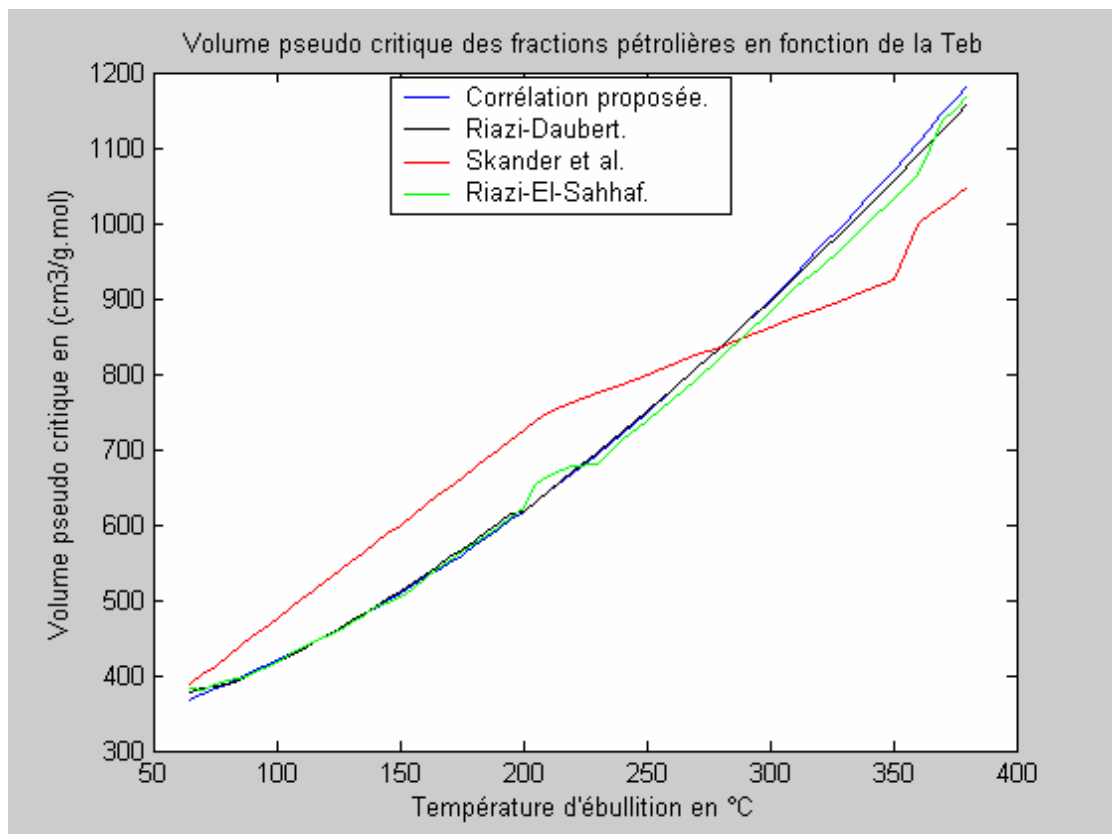
$V_{pc2}$  : Corrélacion de Riazi-Sahhaf.

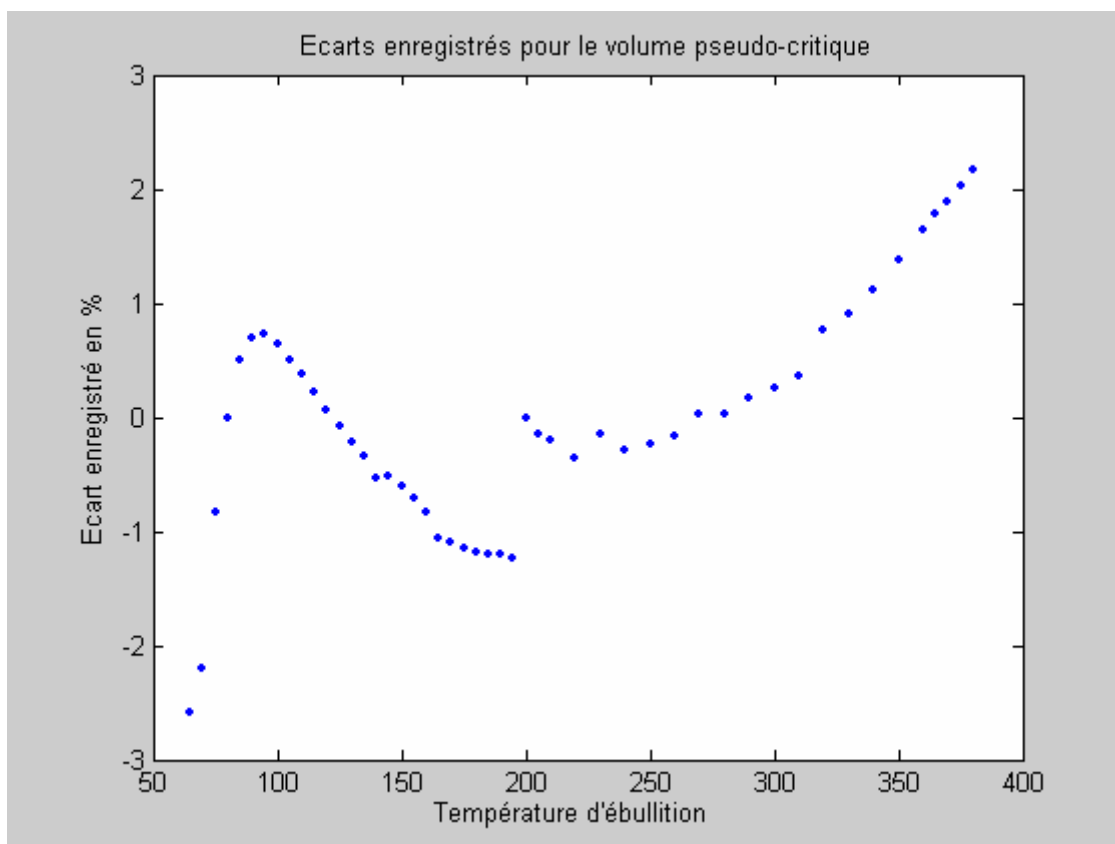
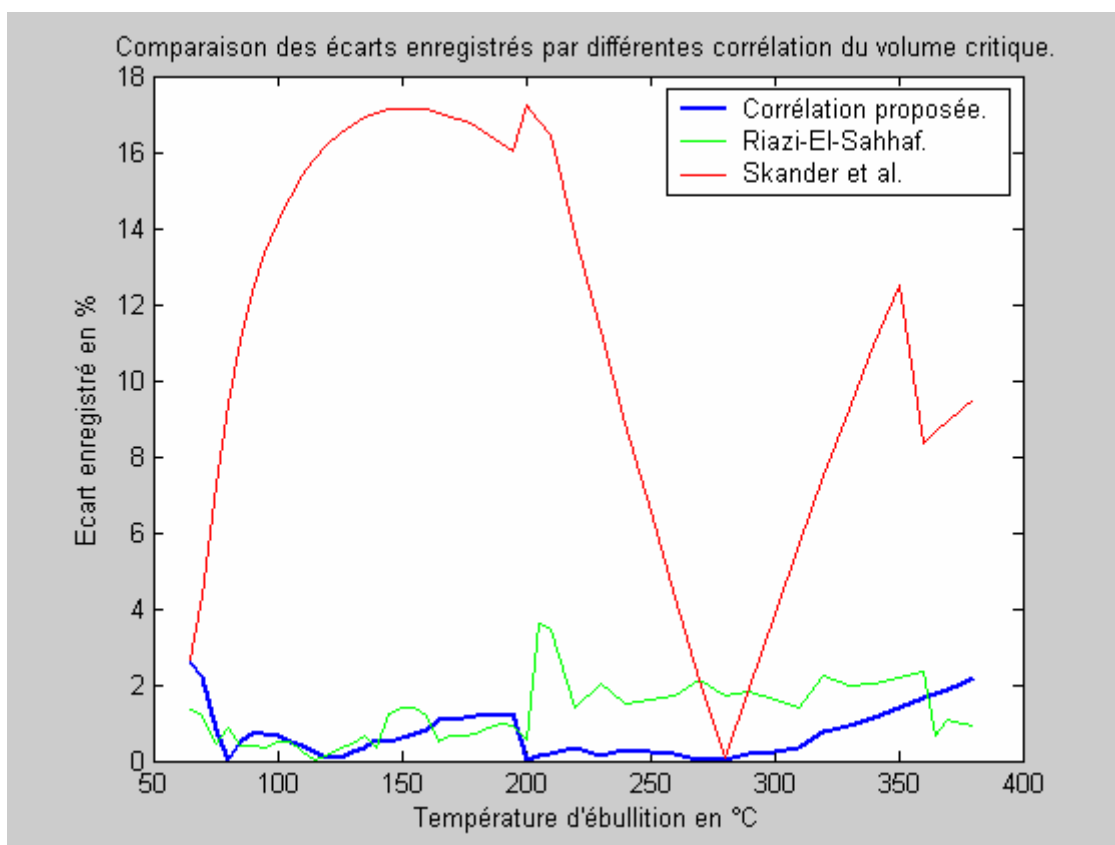
$V_{pc3}$  : Corrélacion de Skander et al.

### V.2-3-2 Etude graphique :

Nous avons représenté dans les figures suivantes les valeurs de la banque de donnée et celle trouvées par le modèle proposé par les algorithmes génétiques et les autres corrélations.

Nous avons aussi représenté l'erreur pour chaque corrélation de chaque valeur de  $V_{pc}$







### V.2-4 Commentaires :

Les propriétés critiques des fractions pétrolières sont calculées en utilisant différentes corrélations. Pour la température critique, les résultats donnés par les corrélations de Lee-Kesler, notre Corrélation, Riazi-Daubert, ont le même ordre de grandeur. La corrélation de Skander et al est elle aussi adéquate pour l'estimation de cette propriété avec un écart moyen de 2.31 % par rapport à la même référence.

Pour l'estimation de la pression critique, nous avons obtenus des valeurs moins bonnes que pour la température critique mais les résultats restent très acceptables puisque l'erreur maximum ne dépasse pas les 5%.

D'après notre étude comparative, nous pouvons dire que la corrélation de Skander et al n'est pas applicable aux coupes lourdes.

Le volume critique lui peut être évalué d'une façon précise par les méthodes de Riazi-Daubert, et la corrélation proposée.

### V.2- 5-1 Facteur acentrique :

L'erreur maximum qu'on peut enregistrer pour le facteur acentrique est de **8 %** et ceci après un grand nombre d'essais d'exécution du programme : le programme converge après **53** générations et un temps d'exécution inférieur à **6** minutes, l'erreur maximale obtenue est de **7.0545**, l'erreur moyenne est de **3.5536** et enfin l'erreur minimale est **0.3136**

Le modèle obtenu par les algorithmes génétiques est le suivant :

$$\omega = 0.002 \times T_{\text{max}} (\text{°C}) \times d_{20} + 0.1847$$

L'application du programme proposé consiste à donner la température d'ébullition et la densité de la fraction choisie pour obtenir directement ce facteur acentrique.

Les résultats obtenus et les valeurs calculées à partir des autres corrélations sont représentés sur le tableau V-4

$\omega$	$\omega_1$	$e_i$	$\omega_{\text{wana}}$	$e_i$	$\omega_{\text{lee}}$	$e_i$
0 . 2 9	0.2693	7.0545	0.2400	17.2414	0.2800	3.4483
0 . 3 0	0.2788	7.0780	0.2400	20.0000	0.2800	6.6667
0 . 3 0	0.2863	4.5533	0.2500	16.6667	0.2900	3.3333
0 . 3 1	0.2945	4.9923	0.2500	19.3548	0.3000	3.2258
0 . 3 1	0.3033	2.1584	0.2600	16.1290	0.3000	3.2258
0 . 3 2	0.3116	2.6306	0.2600	18.7500	0.3100	3.1250
0 . 3 3	0.3198	3.0767	0.2700	18.1818	0.3200	3.0303
0 . 3 3	0.3279	0.6364	0.2800	15.1515	0.3300	0

0 . 3 4	0.3360	1.1626	0.2900	14.7059	0.3300	2.9412
0 . 3 4	0 . 3 4 4 6	1 . 3 3 8 8	0.3000	11.7647	0.3400	0
0 . 3 5	0.3531	0.8809	0.3100	11.4286	0.3500	0
0 . 3 6	0.3616	0.4389	0.3200	11.1111	0.3600	0
0 . 3 6	0.3702	2.8472	0.3200	11.1111	0.3700	2.7778
0 . 3 7	0.3788	2.3827	0.3300	10.8108	0.3700	0
0 . 3 8	0.3876	1 . 9 8 7 1	0.3400	10.5263	0.3800	0
0 . 3 9	0.3955	1.4133	0.3500	10.2564	0.3900	0
0 . 3 9	0.4056	3.9982	0.3500	10.2564	0.4000	2.5641
0 . 4 0	0.4144	3.6025	0.3600	10.0000	0.4100	2.5000
0 . 4 0	0.4232	5.7880	0.3700	7.5000	0.4200	5.0000
0 . 4 1	0 . 4 3 1 4	5 . 2 2 4 4	0.3800	7.3171	0.4300	4.8780
0 . 4 2	0.4386	4.4290	0.3900	7.1429	0.4400	4.7619
0 . 4 3	0.4476	4.0902	0.4000	6.9767	0.4500	4.6512
0 . 4 4	0.4564	3.7205	0.4100	6.8182	0.4600	4.5455
0 . 4 5	0.4652	3.3884	0.4200	6.6667	0.4700	4.4444
0 . 4 6	0.4745	3.1487	0.4200	8.6957	0.4800	4.3478
0 . 4 6 9 6	0.4836	2.9830	0.4300	8.4327	0.4900	4.3441
0 . 4 7 9 0	0.4923	2.7752	0.4400	8.1420	0.5000	4.3841
0 . 4 8 9 2	0.5008	2.3671	0.4500	8.0131	0.5100	4.2518
0 . 4 9 8 5	0.5097	2.2399	0.4600	7.7232	0.5200	4.3129
0 . 5 0 7 5	0 . 5 1 8 8	2 . 2 2 8 6	0.5700	12.3153	0.5300	4.4335
0 . 5 1 7 2	0.5357	3.5831	0.4700	9.1261	0.5400	4.4084
0 . 5 2 5 5	0.5575	6.0951	0.4700	10.5614	0.5400	2.7593
0 . 5 3 9 9	0.5738	6.2856	0.4900	9.2425	0.5700	5.5751
0 . 5 5 9 5	0.5925	5.8981	0.5100	8.8472	0.5900	5.4513
0 . 5 7 9 6	0.6113	5 . 4 7 0 7	0.5300	8.5576	0.6100	5.2450
0 . 5 9 7 0	0.6316	5.7873	0.5400	9.5477	0.6300	5.5276
0 . 6 2 3 5	0.6485	4.0083	0.5700	8.5806	0.6600	5.8540
0 . 6 4 5 3	0.6677	3.4660	0.5900	8.5697	0.6800	5.3773
0 . 6 7 1 5	0.6855	2.0879	0.6200	7.6694	0.7100	5.7334
0 . 6 9 9 0	0 . 7 0 3 3	0.6106	0.6500	7.0100	0.7400	5.8655
0 . 7 1 2 0	0.7268	2.0758	0.6700	5.8989	0.7500	5.3371
0 . 7 4 1 0	0.7446	0.4918	0.7100	4.1835	0.7800	5.2632

0.7666	0.7642	0.3136	0.7500	2.1654	0.8100	5.6614
0.7913	0.7844	0.8732	0.7900	0.1643	0.8300	4.8907
0.8163	0.8048	1.4044	0.8400	2.9033	0.8600	5.3534
0.8472	0.8150	3.8029	0.9000	6.2323	0.8800	3.8716
0.8809	0.8246	6.3937	0.9600	8.9795	0.9100	3.3034
0.8950	0.8347	6.7346	0.9900	10.6145	0.9300	3.9106
0.9088	0.8451	7.0132	1.0300	13.3363	0.9400	3.4331

$\omega$  : La base de donnée (corrélacion de Riazi-Daubert).

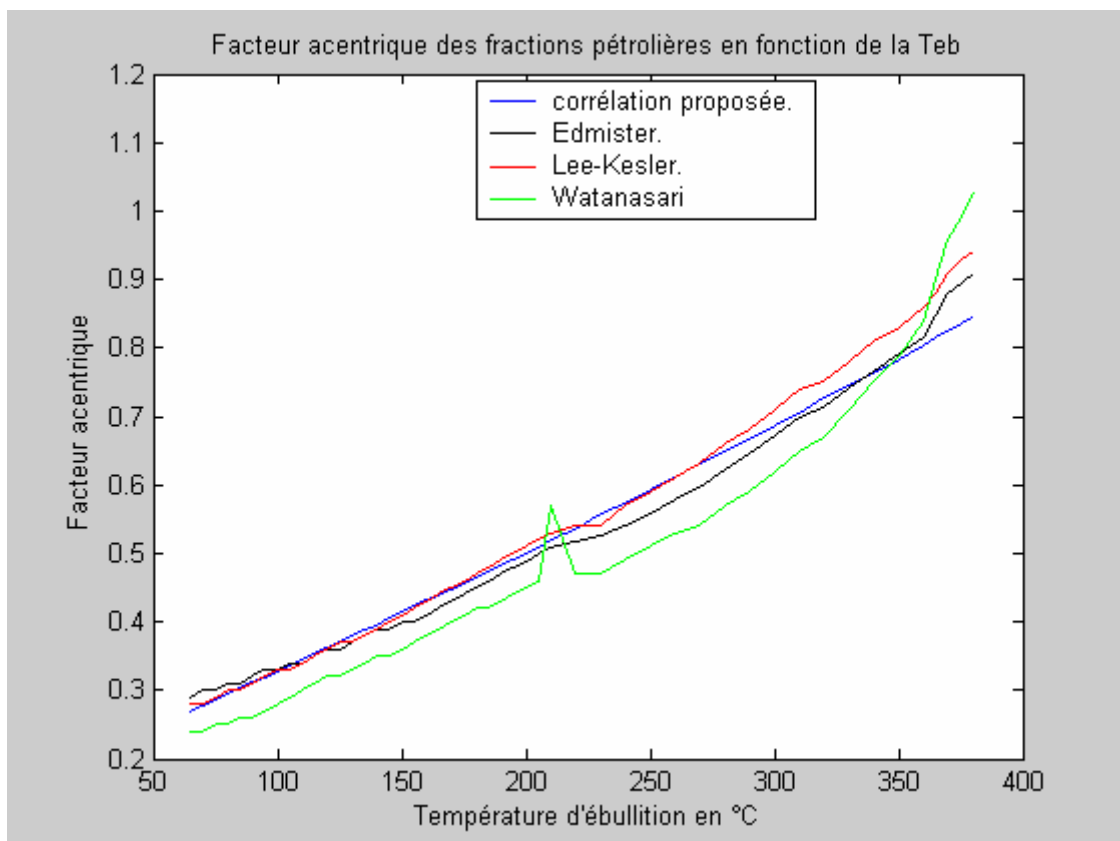
$\omega_1$  : Corrélacion proposée.  $e_i$  : Erreur par rapport à la corrélacion de Riazi-Daubert

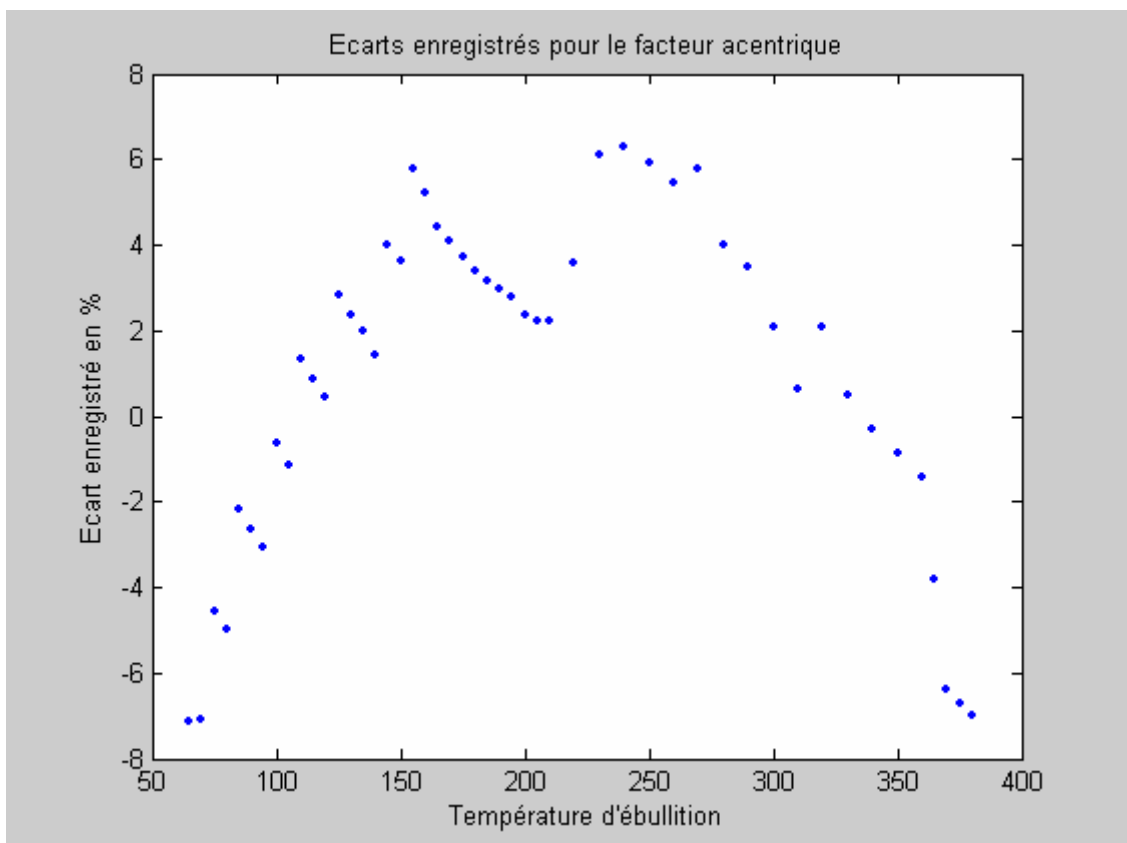
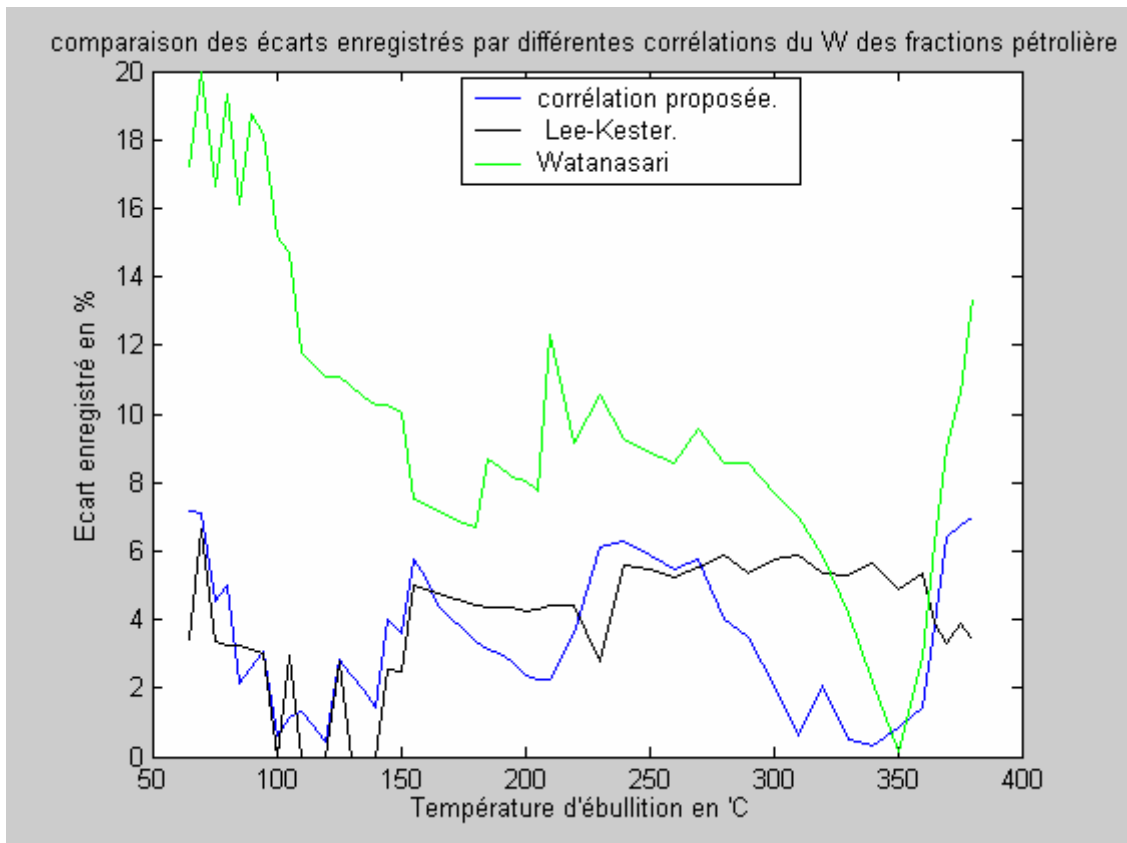
$\omega_{lee}$  : Corrélacion de Lee-kesler.

$\omega_{wana}$  : Corrélacion de Watanasari.

### V.2-5-2 Etude graphique :

Nous avons représenté dans les figures suivantes les valeurs de la banque de données et celle trouvées par le modèle proposé par les algorithmes génétiques et les autres corrélacions. Nous avons aussi représenté l'erreur pour chaque corrélacion de chaque valeur de  $\omega$





### V.2-6 Commentaire :

Ce paramètre a été calculé par différentes corrélations. Les résultats obtenus montrent que l'écart entre les valeurs du facteur acentrique est stable et assez faible pour toutes les corrélations à l'exception de celle de Watanasari.

#### Remarque :

Il est possible que le programme ne converge avec le même nombre de génération et pour la même erreur maximale désirée parce que pour le même espace de recherche les différents opérateurs de l'algorithme génétique fonctionnent d'une façon aléatoire.

### V-3 Conclusion :

Nous avons proposé des modèles d'équations pour chaque propriété, et nous avons utilisé les algorithmes génétiques pour déterminer les coefficients de chaque modèle donnant l'écart maximum le plus faible possible pour chaque propriété étudiée. Les résultats obtenus sont tout à fait acceptables.

Dans le cas de la température critique et le volume critique, les valeurs obtenues se comparent favorablement aux données de référence, puisque l'écart maximum ne dépassent pas 1.25 % pour la  $T_{pc}$  et 2.5 % pour le  $V_{pc}$ .

Pour la pression critique et le facteur acentrique, nous constatons des écarts un peu plus élevés mais restent acceptables. Les résultats peuvent être améliorés soit en changeant l'espace de recherche de chaque paramètre, soit par la proposition d'un autre modèle d'équation ou encore par l'utilisation d'un outil informatique de configuration plus puissante.

Le but visé, qui est l'application de l'intelligence artificielle au calcul des propriétés des fractions pétrolières avec une erreur faible, est atteint.

# Conclusion générale

## Conclusion générale

La connaissance des propriétés physico-chimiques des fractions pétrolières devient nécessaire et fait appel à différentes techniques de caractérisation. Cependant, dans la majeure partie des cas, les méthodes utilisées sont longues et nécessitent un matériel coûteux. C'est pour cela que les techniques, qui permettent d'accéder à ces propriétés, sont très souvent remplacées par des méthodes de calcul, par des corrélations et même par des approches graphiques, dans le but de satisfaire pleinement les objectifs de l'ingénieur du raffinage et de la pétrochimie qui doit désormais connaître, avec davantage de précisions et de détails qu'autrefois, la qualité de chaque produit utilisé dans ses unités.

Dans un tel contexte, nous avons proposé des nouvelles corrélations pour l'estimation de quatre caractéristiques telles que les trois propriétés critiques (température, pression et volume critiques) et le facteur acentrique des fractions pétrolières. A cet effet, plusieurs études ont été menées en parallèle. Il s'agit de :

- Constituer une banque de données (valeurs expérimentales et calculées à partir de corrélations les plus recommandés dans la littérature) nécessaire au développement du modèle de calcul des propriétés étudiées.
- Etablir des corrélations simples pour caractériser les différentes coupes pétrolières à partir de deux paramètres facilement accessibles par expérience telles que la température moyenne d'ébullition et la densité.
- Elaborer un logiciel de calcul des coefficients des modèles proposés en utilisant les Algorithmes Génétiques

Cette nouvelle méthode de caractérisation des pétroles et des coupes qui en découlent a donné des résultats très intéressants pour une utilisation optimum de ces produits.

L'objectif principal donc de ce travail est l'estimation des propriétés critiques et du facteur acentrique des fractions pétrolières, qui sont des mélanges complexes de plusieurs centaines de constituants. La détermination de leur composition fait appel à des expériences longues, coûteuses et difficiles à effectuer.

Il convient de signaler que cette méthode (AG) est applicable dans d'autres domaines par exemple : calcul des échangeurs de chaleur, calcul des réacteurs ainsi que la détermination d'autres propriétés physico-chimiques des hydrocarbures et de leurs mélanges.

## Références bibliographiques.

1. pétrole et gaz arabe. Paris, février 1998.
2. Opec Bulletin, janvier 1998.
3. **P.Wuithier**. Raffinage et Génie Chimique, Tome1, Edition Technip, Paris (1970).
4. Méthodes d'essai des produits pétrolières. Edition Technip (1964).
5. **C.E. Chitour**. Raffinage du pétrole, Tome 1, Edition OPU, Alger (1982).
6. **J-P Wauquier**. Pétrole brut, produits pétroliers. Schéma de fabrication, Tome 1, Edition Technip (1994).
7. **N.Skander**. contribution à la détermination des propriétés physiques des mélanges d'hydrocarbures et des fractions pétrolières par des méthodes du contribution de groupes. Thèse de magister, département Génie Chimique, Ecole Nationale polytechnique, (1997).
8. **M.R.Riazi et T.E.Daubert**, ind. Eng. Chem process, des dev 25-1009 (1986).
9. **O. Kitous**. détermination des propriétés thermodynamiques des fractions pétrolières, Thèse de magister, Dpt de Génie Chimique, Ecole Nationale polytechnique, (1990).
10. **F. Zemmouri**. Elaboration d'un logiciel permettant le calcule des propriétés physique des fractions pétrolières. Thèse de magister, Dpt Génie Chimique, Ecole Nationale Polytechnique (1995).
11. Genetic algorithms in search, optimization end machine learning. Danide Goldberg the University of Alabama, Addison-wesley. Juine 2005.
12. F. Souahi contribution à l'étude physico-chimique des pétroles bruts et des fractions pétroliers comparaison aux méthodes numérique élaborées. Thèse de Doctorat, département Génie Chimique, Ecole Nationale polytechnique, (1999).

### **[Int]. Cite Internet**

1. [www.eudil.fr](http://www.eudil.fr)
2. [www.ai.univ-paris8.fr/renaud/publications/hthese/](http://www.ai.univ-paris8.fr/renaud/publications/hthese/)
3. [www.pétrole\\_monde.com](http://www.pétrole_monde.com)
4. [www.opec.com](http://www.opec.com)
5. [www.gpetrole.cov.ua](http://www.gpetrole.cov.ua)



# Annexe

**La banque de données utilisée et résultats trouvés pour la température critique :**

$T_{eb}$	$d_{20}$	$T_c$	$T_c \text{ calculé}$	erreurs
65	0.6504	502.56	496.7069	1.1647
70	0.6719	512.89	511.6644	0.2390
75	0.6776	518.85	517.9113	0.1809
80	0.6864	525.66	525.8672	0.0394
85	0.6977	533.15	535.2014	0.3848
90	0.7049	539.50	542.2753	0.5144
95	0.7113	545.64	548.9081	0.5989
100	0.7160	551.29	554.6037	0.6011
105	0.7207	556.94	560.2993	0.6032
110	0.7266	562.91	566.6565	0.6656
115	0.7321	568.76	572.7931	0.7091
120	0.7370	574.44	578.5990	0.7240
125	0.7422	580.19	584.5703	0.7550
130	0.7466	585.72	590.1005	0.7479
135	0.7513	591.31	595.7961	0.7587
140	0.7529	596.02	599.7827	0.6313
145	0.7617	602.77	607.7386	0.8243
150	0.7657	608.15	613.0483	0.8054
155	0.7692	613.37	618.0824	0.7683
160	0.7710	618.10	622.1792	0.6600
165	0.7694	621.82	624.4016	0.4152
170	0.7732	627.10	629.6011	0.3988
175	0.7762	632.14	634.3595	0.3511
180	0.7793	637.19	639.1730	0.3112

185	0.7832	642.47	644.4276	0.3047
190	0.7866	647.60	649.4065	0.2789
195	0.7887	652.33	653.6687	0.2052
200	0.7902	656.87	657.6002	0.1112
205	0.7926	661.67	662.0278	0.0541
210	0.7955	666.61	666.7310	0.0182
220	0.7978	671.36	674.2080	0.4242
230	0.8105	679.25	687.4185	1.2026
240	0.8107	687.38	693.7378	0.9249
250	0.8156	696.90	702.6482	0.8248
260	0.8204	706.35	711.5034	0.7296
270	0.8275	716.48	721.6267	0.7183
280	0.8282	724.58	728.2216	0.5026
290	0.8327	733.83	736.9114	0.4199
300	0.8347	742.25	744.2230	0.2658
310	0.8364	750.53	751.3692	0.1118
320	0.8470	761.66	763.4220	0.2313
330	0.8484	769.77	770.4028	0.0822
340	0.8522	778.64	778.7068	0.0086
350	0.8567	787.70	787.3966	0.0385
360	0.8613	796.77	796.1416	0.0789
365	0.8634	804.97	800.4038	0.5672
370	0.8647	812.87	804.2250	1.0635
375	0.8667	817.27	808.4321	1.0814
380	0.8689	821.72	812.7495	1.0917

**La banque de données utilisée et résultats trouvés pour la pression critique :**

$T_{eb}$	$d_{20}$	$P_c$	$P_c$ calculé	erreur
65	0.6504	28.92	30.3823	4.9789
70	0.6719	30.15	29.7574	1.3023
75	0.6776	29.73	29.4263	1.0214
80	0.6864	29.73	29.0377	2.3288
85	0.6977	29.80	28.6025	4.0185
90	0.7049	29.56	28.2436	4.4535
95	0.7113	29.25	27.8995	4.6170
100	0.7160	28.78	27.5871	4.1449
105	0.7207	28.34	27.2747	3.7591
110	0.7266	28.02	26.9399	3.8546
115	0.7321	27.67	26.6126	3.8213
120	0.7370	27.28	26.2965	3.6052
125	0.7422	26.93	25.9748	3.5470
130	0.7466	26.53	25.6679	3.2494
135	0.7513	26.16	25.3555	3.0752
140	0.7529	25.56	25.1008	1.7967
145	0.7617	25.54	24.7121	3.2417
150	0.7657	25.15	24.4127	2.9317
155	0.7692	24.73	24.1226	2.4562
160	0.7710	24.21	23.8641	1.4287
165	0.7694	23.46	23.6689	0.8903
170	0.7732	23.12	23.3732	1.0951
175	0.7762	22.73	23.0924	1.5943
180	0.7793	22.36	22.8097	2.0113
185	0.7832	22.05	22.5122	2.0961
190	0.7866	21.72	22.2239	2.3202
195	0.7887	21.32	21.9599	3.0013

200	0.7902	20.89	21.7070	3.9109
205	0.7926	20.54	21.4373	4.3687
210	0.7955	20.22	21.1584	4.6409
220	0.7978	19.88	20.6656	3.9518
230	0.8105	20.14	19.9794	0.7974
240	0.8107	19.24	19.5257	1.4848
250	0.8156	18.64	18.9845	1.8484
260	0.8204	18.07	18.4453	2.0767
270	0.8275	17.64	17.8632	1.2653
280	0.8282	16.93	17.4002	2.7772
290	0.8327	16.44	16.8665	2.5942
300	0.8347	15.86	16.3793	3.2741
310	0.8364	15.30	15.8977	3.9063
320	0.8470	15.14	15.2505	0.7299
330	0.8484	14.61	14.7745	1.1257
340	0.8522	14.20	14.2538	0.3787
350	0.8567	13.84	13.7201	0.8665
360	0.8613	13.50	13.1845	2.3369
365	0.8634	13.08	12.9205	1.2197
370	0.8647	12.66	12.6713	0.0891
375	0.8667	12.50	12.4091	0.7274
380	0.8689	12.36	12.1432	1.7544

**La banque de données utilisée et résultats trouvés pour le volume critique :**

$T_{eb}$	$d_{20}$	$V_c$	$V_c$ calculé	erreur
65	0.6504	377.98	368.2	2.5523
70	0.6719	383.90	375.5	2.1994
75	0.6776	385.16	381.9	0.8391
80	0.6864	388.94	388.9	0.0131
85	0.6977	394.39	396.3	0.4931
90	0.7049	400.97	403.7	0.6857
95	0.7113	408.35	411.3	0.7200
100	0.7160	416.31	418.9	0.6333
105	0.7207	424.72	426.8	0.4990
110	0.7266	433.49	435.1	0.3709
115	0.7321	442.56	443.6	0.2257
120	0.7370	451.88	452.2	0.0693
125	0.7422	461.43	461.1	0.0709
130	0.7466	471.18	470.2	0.2184
135	0.7513	481.12	479.5	0.3425
140	0.7529	491.25	488.6	0.5387
145	0.7617	501.56	499.0	0.5173
150	0.7657	512.04	508.9	0.6138
155	0.7692	522.68	519.0	0.7076
160	0.7710	533.49	529.0	0.8372
165	0.7694	544.47	538.7	1.0555
170	0.7732	555.61	549.5	1.1002
175	0.7762	566.90	560.4	1.1533
180	0.7793	578.36	571.5	1.1925
185	0.7832	589.98	582.9	1.1949
190	0.7866	601.75	594.5	1.1994
195	0.7887	613.68	606.1	1.2337

200	0.7902	617.84	617.8	0.0099
205	0.7926	630.79	629.8	0.1515
210	0.7955	643.47	642.2	0.1964
220	0.7978	669.27	666.8	0.3616
230	0.8105	695.64	694.6	0.1429
240	0.8107	722.58	720.5	0.2868
250	0.8156	750.10	748.3	0.2419
260	0.8204	778.18	776.9	0.1651
270	0.8275	806.82	807.0	0.0183
280	0.8282	836.03	836.2	0.0181
290	0.8327	865.79	867.2	0.1686
300	0.8347	896.11	898.4	0.2590
310	0.8364	926.98	930.3	0.3603
320	0.8470	958.40	965.8	0.7718
330	0.8484	990.36	999.3	0.8989
340	0.8522	1022.88	1034.3	1.1181
350	0.8567	1055.94	1070.4	1.3738
360	0.8613	1089.54	1175	1.6446
365	0.8634	1106.54	1126.2	1.7779
370	0.8647	1123.68	1144.9	1.8872
375	0.8667	1140.95	1164.0	2.0220
380	0.8689	1158.36	1183.4	2.1650

**La banque de données utilisée et résultats trouvés pour la facteur acentrique :**

$T_{eb}$	$d_{20}$	$W$	$W_{calcul}$ $W_{calculé}$	<i>erreur</i>
65	0.6504	0.29	0.2693	7.0545
70	0.6719	0.30	0.2788	7.0780
75	0.6776	0.30	0.2863	4.5533
80	0.6864	0.31	0.2945	4.9923
85	0.6977	0.31	0.3033	2.1584
90	0.7049	0.32	0.3116	2.6306
95	0.7113	0.33	0.3198	3.0767
100	0.7160	0.33	0.3279	0.6364
105	0.7207	0.34	0.3360	1.1626
110	0.7266	0.34	0.3446	1.3388
115	0.7321	0.35	0.3531	0.8809
120	0.7370	0.36	0.3616	0.4389
125	0.7422	0.36	0.3702	2.8472
130	0.7466	0.37	0.3788	2.3827
135	0.7513	0.38	0.3876	1.9871
140	0.7529	0.39	0.3955	1.4133
145	0.7617	0.39	0.4056	3.9982
150	0.7657	0.40	0.4144	3.6025
155	0.7692	0.40	0.4232	5.7880
160	0.7710	0.41	0.4314	5.2244
165	0.7694	0.42	0.4386	4.4290
170	0.7732	0.43	0.4476	4.0902
175	0.7762	0.44	0.4564	3.7205
180	0.7793	0.45	0.4652	3.3884
185	0.7832	0.46	0.4745	3.1487
190	0.7866	0.4696	0.4836	2.9830
195	0.7887	0.4790	0.4923	2.7752



200	0.7902	0.4892	0.5008	2.3671
205	0.7926	0.4985	0.5097	2.2399
210	0.7955	0.5075	0.5188	2.2286
220	0.7978	0.5172	0.5357	3.5831
230	0.8105	0.5255	0.5575	6.0951
240	0.8107	0.5399	0.5738	6.2856
250	0.8156	0.5595	0.5925	5.8981
260	0.8204	0.5796	0.6113	5.4707
270	0.8275	0.5970	0.6316	5.7873
280	0.8282	0.6235	0.6485	4.0083
290	0.8327	0.6453	0.6677	3.4660
300	0.8347	0.6715	0.6855	2.0879
310	0.8364	0.6990	0.7033	0.6106
320	0.8470	0.7120	0.7268	2.0758
330	0.8484	0.7410	0.7446	0.4918
340	0.8522	0.7666	0.7642	0.3136
350	0.8567	0.7913	0.7844	0.8732
360	0.8613	0.8163	0.8048	1.4044
365	0.8634	0.8472	0.8150	3.8029
370	0.8647	0.8809	0.8246	6.3937
375	0.8667	0.8950	0.8347	6.7346
380	0.8689	0.9088	0.8451	7.0132

## ملخص:

إن الخصائص الفيزيائية و الكيميائية للمنتجات النفطية تتطلب عدة تقنيات لتعيينها. و مع هذا ففي الغالب الطرق المستعملة طويلة و تعتمد على أجهزة عالية الثمن. لهذا فالتقنيات التي تسمح بتعين هذه القياسات تعوض غالبا بمناهج حسابية و بعلاقات خاصة

إن العمل المنجز يتمثل في استعمال الذكاء الاصطناعي (الخوارزميات الوراثية) لتمييز كل الأجزاء النفطية التي هي عبارة عن خلائط معقدة نتيجة لمكوناتها المتعددة.

إن جميع العلاقات المقترحة في هذه الدراسة قد أعطت نتائج جد هامة و سمحت بتثمين معتبر للعينات النفطية زيادة على ربح الوقت و المال

الأجزاء النفطية, الخوارزميات الوراثية, الذكاء الاصطناعي, الخصائص الفيزيائية

## Abstract:

Different characterization techniques must be in order to determine physical and thermodynamically properties of hydrocarbon products. However, in most cases these techniques are lengthy and require costly equipment. Thurs, other methods based on numerical and graphical approaches are favored.

Our study consists to determine the critical properties of petroleum fractions using the Genetic Algorithms

The proposed correlations we to obtain interesting result with cost and experimental time which have been significant reduced.

**Mots clés: Thermodynamically, critical properties, petroleum fractions, Genetic Algorithms**

## Résumé :

La connaissance des propriétés physiques et thermodynamiques des produits pétrolières fait appel à différentes techniques de caractérisation. Cependant, dans la majeure partie des cas, les méthodes utilisées sont longues et dépendant d'un matériel coûteux. C'est pourquoi les techniques permettant la détermination de ces grandeurs sont souvent remplacées par des méthodes de calcul, par des corrélations

Notre travail consiste à développer une nouvelle méthode basée sur les Algorithmes Génétiques pour estimer les propriétés critiques et le facteur acentrique des fractions pétrolières.

Les différentes corrélations proposées ont donné des résultats intéressants.

**Mots clés : thermodynamiques, propriétés critiques, facteur acentrique, fractions pétrolières, Algorithmes Génétiques, corrélations.**