

P0012/05B  
Ecole Nationale Polytechnique

Département de Génie Chimique

المدرسة الوطنية المتعددة التقنيات  
المكتبة — BIBLIOTHEQUE  
Ecole Nationale Polytechnique



Projet de fin d'études  
En vue de l'obtention de diplôme d'ingénieur d'état

Thème :

Détermination des propriétés critiques des  
hydrocarbures par le développement d'un  
nouveau modèle mathématique basé sur les  
Algorithmes Génétiques

Etudié par :

**MERABET AYECHÉ Fouzi**  
**TIDJANI Hamza**

Proposé et dirigé par :

**Dr F.SOUAHI**

## ملخص:

إن معرفة الخصائص الفيزيائية و الترمودينامكية للمحروقات تستدعي عدة تقنيات مختلفة؛ إن طرق التخصيص التجريبية معروفة بطولها و بثمنها الباهض, مما أدى إلي استبدالها بطرق حسابية.

يتمثل عملنا في تقديم تقنية جديدة أساسها الخوارزميات ال وراثية من أجل إيجاد الخصائص الحرجة للأجسام النقية. النموذج المقترح أعطي نتائج مقبولة

## الكلمات المفتاحية:

الخصائص الفيزيائية و الترمودينامكية  
الخوارزميات ال وراثية  
الخصائص الحرجة  
الأجسام النقية

## Résumé :

La connaissance des propriétés physique et thermodynamique des hydrocarbures fait appel à différentes techniques de caractérisation. Ces méthodes expérimentales sont très souvent longue et dépendent d'un matériel coûteux. Elles sont souvent remplacées par des méthodes de calcul.

Notre étude consiste à développer une nouvelle méthode basée sur les Algorithme Génétique pour estimer les propriétés critiques des corps purs. Le modèle proposé donne des résultats acceptables.

## Mots clés :

propriétés physique et thermodynamique  
les propriétés critiques  
les Algorithme Génétique  
corps purs

## Abstract :

Different characterization techniques must be used in order to determine physical and thermo dynamical properties of hydrocarbons. These experimental methods are lengthy and require courtly equipments. Thus, numerical approaches are favored.

Our study consists to determine the critical properties of pure components using the Genetic Algorithms; the obtained results are good.

## Key words:

physical and thermo dynamical properties  
critical properties  
Genetic Algorithms  
pure components

## *Remerciements*

*Ce travail fait partie des activités de recherche du Laboratoire de Valorisation des Energies Fossiles dirigé par Professeur CHITOUR que nous remercions.*

*Nous tenons à exprimer nos vifs remerciements à tous nos professeurs qui ont contribué à notre formation en particulier notre promotrice Mme Souahi et son fils Abdeldjallil pour leurs conseils et l'aide qu'ils nous ont apporté .*

*Nous voulons aussi exprimer toute notre gratitude et nos remerciements à tout le personnel du département Génie Chimique.*

*Nos remerciements vont aussi à M Ahmed Zaid chef du département de Génie Chimique ; pour avoir consacré un peu de son précieux temps pour les nombreux services rendus.*

## *Dédicaces*

*Louanges à Dieu, m'ayant comblé de ses bienfaits, j'ai pu réaliser  
cette thèse, et paix et bénédiction sur le prophète Mohamed, sa  
famille, ainsi que tous ceux qui ont emprunté sa voix.  
A ceux qui ont beaucoup soufferts pour que je serai bien, à mes  
chères parents.*

*A mes frères et sœurs et particulièrement à la nouvelle née Sabrina que  
je n'ai pas encore vue.*

*A mes fidèles amis : Abdallah, Bilel & Youcef*

*A mon cher binôme Hamza qui m'a soutenu dans les moments difficiles*

*...Je dédie cette thèse.*

*Fouzi*

## *Dédicaces*

*Je dédie ce modeste travail à ceux qui m'ont donné tout dont j'avais besoin, à mes parents.*

*A mes frères et ma sœur unique.*

*A mes oncles et tantes et tout particulièrement à TIDJANI Abdelkarim.*

*A mes cousins et cousines.*

*A tous mes Amis de la résidence universitaire BOURAOUI Amar et particulièrement à BENAÏSSA Smain.*

*A la mémoire du petit Smain*

# SOMMAIRE

	Page
Introduction générale .....	1
Chapitre I : Méthodes de détermination des propriétés critiques	
I.1. Fondement thermodynamique.....	2
I.2. Mesures expérimentales des paramètres critiques.....	3
I.3. Estimations des propriétés critiques.....	4
I.3.1. Température critique.....	4
I.3.1.1. Estimation de la température critique des corps purs.....	4
I.3.2.2. Estimation de la température critique des mélanges de corps purs.....	5
I.3.3.3. Estimation de la température critique des fractions pétrolières.....	5
I.3.2. Pression critique.....	6
I.3.2.1. Estimation de la pression critique des corps purs.....	6
I.3.2.2. Estimation des pression critiques des mélanges de corps purs.....	7
I.3.2.3. Estimation des pressions critiques des fractions pétrolières.....	7
I.3.3. Volume critique.....	8
I.3.3.1. Estimation des volumes critiques des corps purs.....	8
I.3.3.2. Estimation des volumes critiques des mélanges de corps purs.....	9
I.3.3.3. Estimation des volumes critiques des fractions pétrolières.....	9
Chapitre II : Introduction aux Algorithmes Génétiques (AG)	
II.1. Historique et définition de la théorie de Darwin.....	11
II.2. Principe de fonctionnement des algorithmes génétiques.....	12
II.3. Codage.....	13
II.3.1 Codage binaire.....	14
II.3.2 Codage réel.....	16
II.4. Définition de la fonction fitness.....	16
II.5. Opérateur de sélection.....	17
II.5.1. Définition de la sélection.....	17
II.5.2. Sélection par roulette ( <i>wheel</i> ).....	17

II.5.3. Sélection par rang.....	18
II.5.4. Sélection par steady-state.....	18
II.5.5. Sélection par tournois.....	19
II.6. Opérateur de croisement.....	19
II.6.1. Croisement binaire.....	19
II.6.1.1. Croisement en un point.....	19
II.6.1.2. Croisement en deux points.....	20
II.6.1.3. Croisement uniforme.....	20
II.6.2. Croisement réel.....	21
II.6.2.1. Croisement simple.....	21
II.6.2.2. Croisement arithmétique.....	21
II.6.2.3. Croisement heuristique.....	21
II.6.2.4. Croisement MPX.....	22
II.7. Opérateur de mutation.....	22
II.7.1. Mutation binaire.....	22
II.7.2. Mutation réelle.....	24
II.7.3. Mutation non uniforme.....	24
<b>Chapitre III : Application des AG pour l'élaboration du modèle</b>	
III.1. Enoncé du problème.....	25
III.2. Codage réel des paramètres d'entrées.....	26
III.3. Création de la population initiale.....	26
III.4. Calcul du fitness.....	27
III.5. Opérateur de sélection.....	29
III.6. Opérateur de croisement.....	31
III.7. Opérateur de mutation.....	34
III.8. Teste d'arrêt.....	37
<b>Chapitre VI : Résultats et commentaires</b>	
IV.1. Introduction.....	38
IV.2. Résultats obtenus.....	38

IV.2.1. Température critique.....	38
IV.2.2. Pression critique.....	41
IV.2.3. Volume critique.....	44
IV.3. Etude comparative avec le logiciel SigmaPlot.....	46
IV.4. Discussions.....	47

## Conclusion

### Annexe I : Banque de données

- I.1. Température critique
- I.2. Pression critique
- I.3. Volume critique

### Annexe II : Résultats obtenus

- II.1. Température critique
- II.2. Pression critique
- II.3. Volume critique





Terme	Signification
AG	Algorithmes génétiques
$d_{20}$	Densité à 20 °C
Fitness	L'erreur du modèle en %
Individu	Une ligne de la matrice POP
$P_c$	Pression critique (bar)
$p_c$	Probabilité de croisement
$p_m$	Probabilité de mutation
Rand	Un chiffre tiré au hasard entre [0,1]
$T_c$	Température critique (K)
$T_{eb}$	Température d'ébullition (K)
$V_c$	Volume critique ( $Cm^3/mol$ )

### Remarque :

Le micro-ordinateur utilisé pour les calculs est un Pentium 4 de fréquence 2.26 Ghz et d'une mémoire DDRAM de 128 MO. Le logiciel utilisé pour la programmation est le MATLAB version 6.5.

المدرسة الوطنية المتعددة التقنيات  
BIBLIOTHEQUE — المكتبة  
Ecole Nationale Polytechnique

# Introduction Générale



L'énergie, aujourd'hui plus qu'auparavant, constitue une ressource vitale incontournable pour toute l'humanité.

Malgré la nouvelle impulsion donné aux énergies renouvelables, le pétrole est toujours considéré comme la composante essentielle au développement économique, social et technologique des différents pays.

L'exploitation et la valorisation des hydrocarbures passent par une connaissance approfondie de leurs caractéristiques pour une mise en œuvre optimale de ces produits.

L'importance donc que revêt cette identification dans l'industrie du pétrole est considérable. Devant l'impossibilité d'effectuer certaines mesures expérimentales, différentes études ont été réalisées pour déterminer indirectement ces propriétés.

Dans un tel contexte, ce travail s'est fixé comme objectif la mise au point d'un nouveau modèle numérique pour estimer les propriétés critiques difficilement accessibles par expérience.

Il s'agit de l'élaboration d'un outil de calcul basé sur les algorithmes génétiques dont le principe de base s'inspire de la théorie de Darwin sur l'évolution des espèces :

En permettant aux seuls individus bien adaptés à l'environnement de se reproduire, la nature assure la pérennité de leurs meilleures caractéristiques; lesquelles caractéristiques se recombinent entre elles (chaque enfant reçoit de bonnes caractéristiques à la fois de son père et de sa mère) pour former à chaque génération de nouveaux individus toujours mieux adaptés à leur environnement.

# Chapitre I

Méthodes de détermination  
des propriétés critiques.

### I.1. Fondement thermodynamique[1]

Dans un diagramme de coordonnées pression et température (figure I.1) les équilibres de sublimation, fusion et vaporisation d'une substance pure sont représentés par trois courbes qui passent par le point triple où coexistent trois phases en équilibre. La courbe de vaporisation est limitée quant à elle au point critique « c », borne au delà de laquelle la coexistence des deux phases liquide et vapeur n'est plus possible. Ce point qui a pour coordonnées  $T_c$ ,  $P_c$  et  $V_c$  (température, pression et volume critiques) correspond à l'identité des propriétés du liquide et de la vapeur. Plus particulièrement, la chaleur de vaporisation devient nulle en ce point.

Dans le cas d'un mélange, la courbe d'équilibre de vaporisation se décompose en deux courbes : courbe de bulle et courbe de rosée, qui se rejoignent au point critique (figure.I.2). pour une pression et une température égales respectivement à la pression et à la température critiques, les compositions des deux phases deviennent identiques.

Ces propriétés critiques sont des caractéristiques physiques des corps purs. La température et la pression critiques définissent le point limite supérieur de la courbe de tension de vapeur, au-delà duquel le changement de phase n'est plus observé.

La connaissance des propriétés critiques des substances pures, des mélanges de corps purs et des fractions pétrolières est d'une grande importance pour corréler ou estimer les propriétés inaccessibles ou difficilement accessibles expérimentalement.

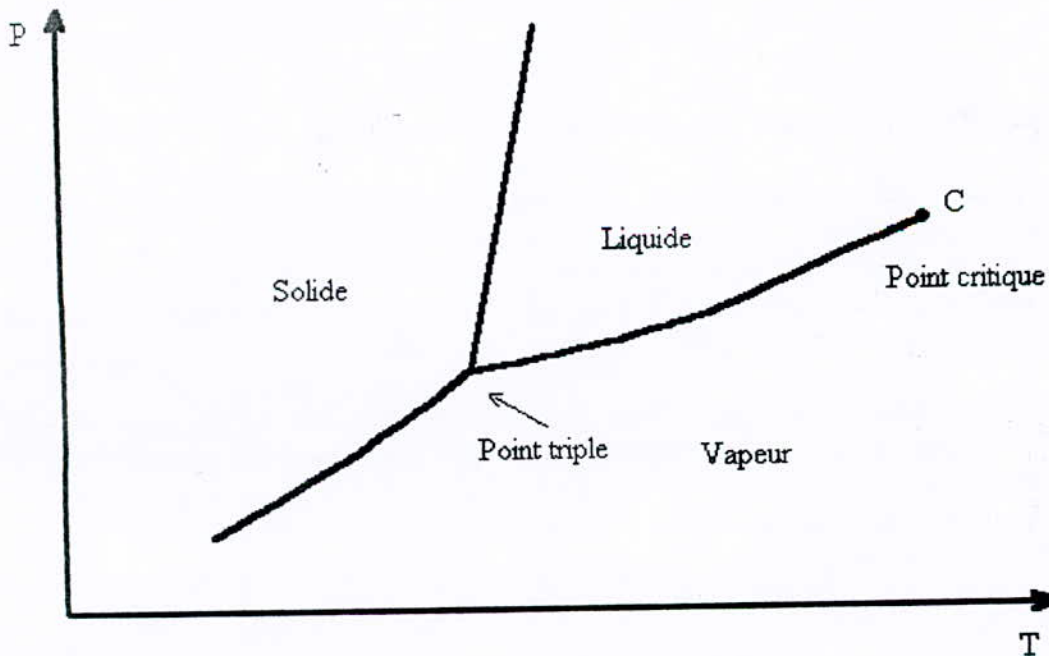


Figure I.1: Diagramme de phases d'un corps pur

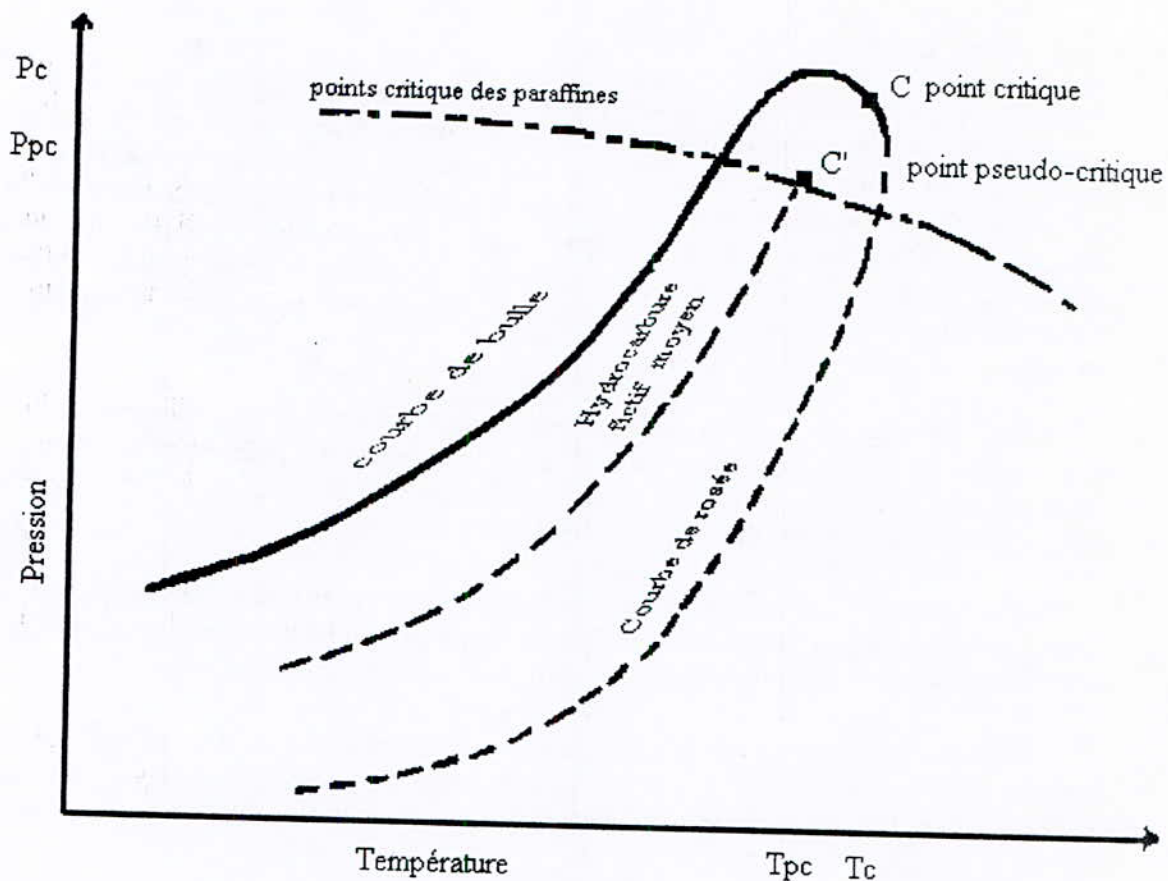


Figure I.2 : Diagramme de phases d'un mélange de corps purs

## I.2. Mesure expérimentale des paramètres critiques [2]

La détermination expérimentale des propriétés critiques nécessite un appareillage assez coûteux et sophistiqué. Le principe de mesure pour un fluide consiste à chauffer celui-ci jusqu'à la disparition du ménisque séparant les phases liquide et vapeur et formation d'un brouillard opalescent. La température, la pression et le volume correspondants à cet état ne sont autres que les propriétés critiques.

Deux méthodes ont été citées dans la littérature : celle dite de l'ampoule (Smith et Teja, 1987) et la méthode de l'écoulement (Rosenthal et Teja, 1989) ces deux méthodes sont dotées d'un dispositif d'enregistrement et de visualisation par vidéo, l'ensemble relié à un ordinateur pour le contrôle de l'opération de mesure.

### I.3. Estimations des propriétés critiques [1 et 2]

Les propriétés critiques sont d'un intérêt certain. Elles sont nécessaires pour appliquer la loi des états correspondants, pour utiliser les équations d'états et pour déterminer les propriétés thermodynamiques. Leur connaissance est indispensable dans les procédés d'extraction liquide dans le domaine supercritique.

#### I.3.1. Température critique

##### I.3.1.1. Estimation de la température critique des corps purs

Les température critiques des corps purs peuvent être générées soit à partir des méthodes de contribution de groupes ce qui exige la connaissance de la structure chimique, soit encore par des méthodes empiriques. Parmi ces méthodes, nous citons celle de Lydersen (1955), Nokay (1959. Ambrose), Klincewicz et Reid (1984), Joback et Reid (1984), Sanchez (1985), constantinou et Al (1994), Riazi et El-Sahhaf (1995), ect...

Nous donnons ci-après les équations des méthodes les plus connues et faisant intervenir la structure chimique des molécules :

Auteurs	Equations	
Lydersen (1955)	$T_c = T_{eb} [0.567 + \sum \delta T - (\sum \delta T)^2]^{-1}$	$\delta T$ : incrément
Joback et Reid (1984)	$T_c = T_{eb} [0.584 + 0.965 \sum \delta T - (\sum \delta T)^2]^{-1}$	
Nokay (1959)	$\log(T_c) = A + B \log(\text{SpGr}) + C \log(T_{eb})$	A,B,C,D Constantes dépendant de la famille
Ambrose (1980)	$T_c = T_{eb} [1 + (1.242 + \sum \delta T) ]$	
Fedors (1982)	$T_c = 535 [\log(m + \sum \delta T) - 0.350]$	m : constante
Constantinous et coll (1994)	$\exp(T_c / 181.128) = \sum \delta T$	

Riazi et El-Sahhaf (1995)	$(T_c / T_{eb})^{-1} = a - \exp(b - c M^d)$	a, b, c et d : constantes dépendant de la famille
Souahi et coll	$T_c = a + b/[1+(N/c)^d]$	N : nombre de carbon a, b, c et d : constantes dépendant de la famille

### I.3.1.2. Estimation de la température critique des mélanges de corps purs

Pour les mélanges de corps purs, il a été remarqué que les coordonnées du point critique vrai, situé à la jonction des courbes de bulle et de rosée, ne vérifient pas les équations des états correspondants. De ce fait, il est apparu nécessaire d'introduire de nouveaux paramètres au même titre que les propriétés critiques des corps purs, afin que la loi des états correspondants puisse être étendue aux mélanges. Ainsi, en établissant une équivalence entre le mélange et un hydrocarbure pur paraffinique dont la courbe de tension de vapeur est comprise entre les courbes de bulle et de rosée du mélange, un nouveau point critique C' a été défini, situé à l'intersection de la courbe de tension de vapeur de l'hydrocarbure fictif et du lieu des points critiques des hydrocarbures paraffiniques (figure.I.2). Ce nouveau point critique C' est appelé point pseudo-critique du mélange et a pour coordonnées :

$T_{Pc}$  : température pseudo-critique

$P_{Pc}$  : pression pseudo-critique

$V_{Pc}$  : volume pseudo-critique

Les propriétés pseudo-critiques des mélanges sont obtenues généralement à partir de la règle de KAY (1936) qui s'exprime comme suit dans le cas de la température pseudo-critique :

$$T_{Pc} = \sum x_i T_{Ci} \quad \text{équation I.1}$$

Dans cette équation,  $x_i$  représente la fraction molaire du constituant i dans le mélange et  $T_{Ci}$  sa température critique.

### I.3.1.3. Estimation de la température critique des fractions pétrolières

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, nous disposons de nombreuses corrélations qui nécessitent la connaissance d'autres paramètres physiques tels que la température moyenne d'ébullition et la specific-gravity. Nous citons quelques unes des méthodes parmi les plus connues.



Auteurs	Equations
Additivité	$T_{Pc} = T_{cP} X_P + T_{cN} X_N + T_{cA} X_A$ <p> <math>X_i</math> : composition  <math>T_{cP}</math> : pression critique des Paraffines  <math>T_{cN}</math> : pression critique des Naphtènes  <math>T_{cA}</math> : pression critique des Aromatiques </p>
Lee et Kesler (1976)	$T_{Pc} = 341.7 + 811 \text{ SpGr} + (0.4244 + 0.1174 \text{ SpGr}) T_{eb} + (0.4669 - 3.2623 \text{ SpGr}) 10^5 / T_{eb}$ <p style="text-align: right;">SpGr : densité</p>
Riazi et Daubert (1980)	$T_{Pc} = 24.2787 T_{eb}^{-2.3125} \text{ SpGr}^{2.3201}$
Watanasiri et coll (1985)	$\ln T_{Pc} = -9.3906 10^{-5} T_{eb} + 0.03095 \ln M + 1.11067 \ln T_{eb} + M(0.078154 \text{ SpGr}^{1/2} - 0.061064 \text{ SpGr}^{1/3} - 0.0169438 \text{ SpGr})$ <p>M : masse molaire</p>
Riazi et Daubert (1987)	$T_{Pc} = a \exp (b \theta_1 + c \theta_2 + d \theta_1 \theta_2) \theta_1^e \theta_2^f$ <p> <math>(\theta_1, \theta_2)</math> couples de propriétés tels que <math>(T_{eb}, \text{SpGr})</math>  a, b, c, d et f : constantes </p>
Les réseaux de neurones	[3,4,5 et 6]

Si la fraction pétrolière contient un nombre fini d'hydrocarbures (cas des fractions légères) et si sa composition qualitative et quantitative est connue, la règle d'additivité de l'équation (I.1) peut être utilisée.

## I.3.2. Pression critique

### I.3.2.1. Estimation de la pression critique des corps purs

Parmi les méthodes qui permettent de prédire les pressions critiques des corps purs, nous donnons ci-dessous les méthodes, considérées comme les plus importantes, permettant de prédire la pression critique.

Auteurs	Equations	
Lydersen (1955)	$P_c = M (\sum \sigma P + 0.34)^{-2}$	$\sigma P$ : incrément M : masse molaire
Joback et Reid (1984)	$P_c = (\sum \sigma P + 0.133 + 0.0032 N)^{-2}$  N : nombre d'atome de carbone dans la molécule $\sigma P$ : incrément	
Ambrose (1980)	$P_c = M (\sum \sigma P + 0.339)^{-2}$	$\sigma P$ : incrément M : masse molaire
Constantinou et Coll (1995)	$(P_c - 1.3705)^{-0.5} = 0.10022 + \sum \sigma P$	$\sigma P$ : incrément
Riazi et El-Sahhaf (1995)	$P_c = a - \exp(b - c M^d)$	a, b, c et d : constantes dépendant de la famille M : masse molaire
Souahi et coll	$P_c = a + b / [1 + (N/c)]$	N : nombre de carbone a, b, c et d constantes dépendant de la famille

### I.3.2.2. Estimation des pression critiques des mélanges de corps purs

La règle de KAY (1936) relative au calcul des pressions pseudo-critiques des mélanges s'écrit :

$$P_{Pc} = \sum x_i P_{ci}$$

$x_i$  et  $P_{ci}$  sont respectivement la fraction molaire et la pression critique du constituant i dans le mélange.

### I.3.2.3. Estimation des pressions critiques des fractions pétrolières

Les pressions critiques des fractions pétrolières peuvent être estimées en utilisant différentes corrélations dont les plus connues sont celle de Lee-Kesler (1976), Riazi-Daubet (1980) et Watanasiri et Coll (1985).

Auteurs	Equations
Additivité	$P_{Pc} = P_{cP} X_P + P_{cN} X_N + P_{cA} X_A$ <p> <math>X_i</math> : composition  <math>P_{cP}</math> : pression critique des Paraffines  <math>P_{cN}</math> : pression critique des Naphtènes  <math>P_{cA}</math> : pression critique des Aromatiques </p>
Lee et Kesler (1976)	$\ln P_{Pc} = 83634 - 0.566 / SpGr - (0.24244 - 2.2898 / SpGr - 0.11857 / SpGr^2) 10^{-3} T_{eb} - (14685 - 3648 / SpGr - 16977 / SpGr^2) 10^{-10} T_{eb}^3$ <p>SpGr : densité</p>
Riazi et Daubert (1980)	$P_{Pc} = 3.12281 T_{eb}^{-2.3125} SpGr^{2.3201}$ <p>SpGr : densité</p>
Watanasiri et Coll (1985)	$\ln P_{Pc} = 3.9543 + 0.70682 (T_c / V_c)^{0.8} - 4.8400 M / T_c - 0.15919 T_{eb} / M$
Riazi et Daubert (1987)	$P_c = a \exp (b \theta_1 + c \theta_2 + d \theta_1 \theta_2) \theta_1^e \theta_2^f$ <p> <math>(\theta_1, \theta_2)</math> couples de propriétés tels que <math>(T_{eb}, SpGr)</math>  <math>a, b, c, e</math> et <math>f</math> : constantes </p>
Les réseaux de neurones	[3, 4, 5 et 6]

### I.3.3. Volume critique

#### I.3.3.1. Estimation des volumes critiques des corps purs

Parmi les méthodes de détermination des volumes critiques des corps purs, nous pouvons citer les méthodes suivantes :

Auteurs	Equations	
Lydersen (1955)	$V_c = 40 + \sum \sigma V$	$\sum \sigma V$ : incrément
Riedel (1954)	$V_c = R T_c / P_c [3.72 + 0.26 (\alpha - 7)]^{-1}$ $\alpha = 0.9076 [1 + \ln P_c / (T_c - 1)]$	
Joback et Reid (1984)	$V_c = 175 + \sum \delta V$	
Constantinou et Coll (1994)	$V_c = 40 + \sum (\sigma V - 0.00435)$	
Riazi et El-Sahhaf (1995)	$(V_c / M)^{-1} = a - \exp(b - c M^d)$	a, b, c et d : constantes dépendant de la famille M : masse molaire
Souahi et coll (1995)	$V_c = a + b N$	N : nombre d'atome de carbone a et b : constantes dépendant de la famille

### I.3.3.2. Estimation des volumes critiques des mélanges de corps purs

Le volume pseudo-critique des mélanges de corps purs s'obtient lui aussi moyennant la règle de KAY (1936) en pondérant les volumes critiques des constituants en moles :

$$V_{pc} = \sum x_i V_{ci}$$

$V_{pc}$  : est le volume pseudo-critique du mélange.  
 $x_i$  et  $V_{ci}$  étant respectivement la fraction molaire et le volume critique du constituant i dans le mélange.

### I.3.3.3. Estimation des volumes critiques des fractions pétrolières

Riazi-Daubert (1980), (1987) ainsi que Watanasri et Al (1985) ont proposé des corrélations pour l'estimation des volumes critiques des fractions pétrolières de même type que celles proposées pour les températures et pression critiques. Les équations qu'ils proposent sont les suivantes.

Auteurs	Equations
Additivité	$V_{Pc} = V_{cP} X_P + V_{cN} X_N + V_{cA} X_A$ <p style="margin-left: 40px;"> <i>X<sub>i</sub></i> : composition  <i>V<sub>cP</sub></i> : pression critique des Paraffines  <i>V<sub>cN</sub></i> : pression critique des Naphtènes  <i>V<sub>cA</sub></i> : pression critique des Aromatiques         </p>
Riazi et Daubert (1980)	$V_{Pc} = 7.5214 \cdot 10^5 T_{eb}^{0.2896} SpGr^{0.7666}$ <p style="margin-left: 40px;">SpGr : densité</p>
Watanasiri et al (1985)	$\ln V_{Pc} = 80.4479 + 129.8083 SpGr + 63.1750 SpGr^2 - 13.1750 SpGr^3 + 1.10108 \ln M + 42.1958 \ln SpGr$ <p style="margin-left: 40px;">SpGr : densité ; M : masse molaire</p>
Riazi et Daubert (1987)	$V_{Pc} = a \exp(b \theta_1 + c \theta_2 + d \theta_1 \theta_2) \theta_1^e \theta_2^f$ <p style="margin-left: 40px;"> <i>(θ<sub>1</sub>, θ<sub>2</sub>)</i> couples de propriétés tels que (<i>T<sub>eb</sub></i>, SpGr)  <i>a, b, c, e</i> et <i>f</i> : constantes         </p>
Les réseaux de neurones	[3, 4, 5 et 6]

# Chapitre II

Introduction aux Algorithmes Génétiques

## II.1. Historique et définition de la théorie de Darwin [7]

Les hommes ont observés le monde depuis des temps immémoriaux. Cette observation les a menés à établir des théories permettant d'expliquer les régularités de la nature. Puis à l'observation, la théorie et l'expérience est venu s'ajouter un nouvel outil scientifique : la modélisation. Celle-ci a pleinement acquis ses lettres de noblesse avec l'avènement de l'informatique. Elle s'est depuis imposée comme un outil indispensable à la pratique de la science, trouvant sa place entre l'observation, la théorie et l'expérience, et interagissant avec elles. Loin d'être un substitut à l'expérience, la simulation informatique se révèle précieuse pour choisir l'expérience qu'il convient de réaliser pour obtenir le résultat voulu. Dans certains domaines, les expériences ont la fâcheuse tendance à devenir de plus en plus coûteuses. La simulation permet donc un gain de temps et une économie substantielle en réduisant le nombre d'expériences à mener. On peut parler désormais d'expérimentation virtuelle.

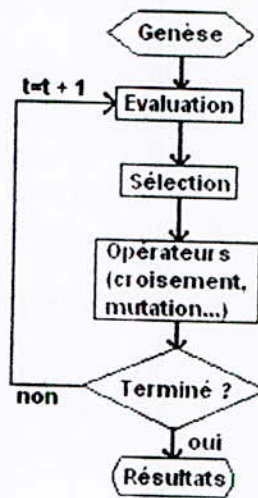
C'est en 1975 que John Langton parle pour la première fois des algorithmes génétiques, une technique de programmation qui s'inspire du principe de l'évolution des espèces décrit par Darwin, cette théorie repose sur deux postulats simples :

- Dans chaque environnement, seul les espèces les mieux adaptées perdurent au cours des temps, les autres étant condamnées à disparaître.
- Au sein de chaque espèce, le renouvellement des populations est essentiellement dû aux meilleurs individus de l'espèce.

Mais ce n'est que récemment (vers le début des années 90) que cette technique arrive sur le devant de la scène, s'inscrivant un peu, avec les techniques liées à la vie artificielle d'une façon plus générale, comme une voie de renouvellement pour le génie logiciel et l'intelligence artificielle (certains préféreront parler d'informatique avancée).

C'est en effet cette légèreté de mise en oeuvre qui fait tout le charme des algorithmes génétiques et qui donne cette impression rafraîchissante d'avoir insufflé un peu de vie et de créativité dans l'ordinateur. Ce seul aspect révolutionnera peut-être la façon que l'on a de voir les machines qui nous entourent.

## II.2. Principe de fonctionnement d'un AG [7] et [8]

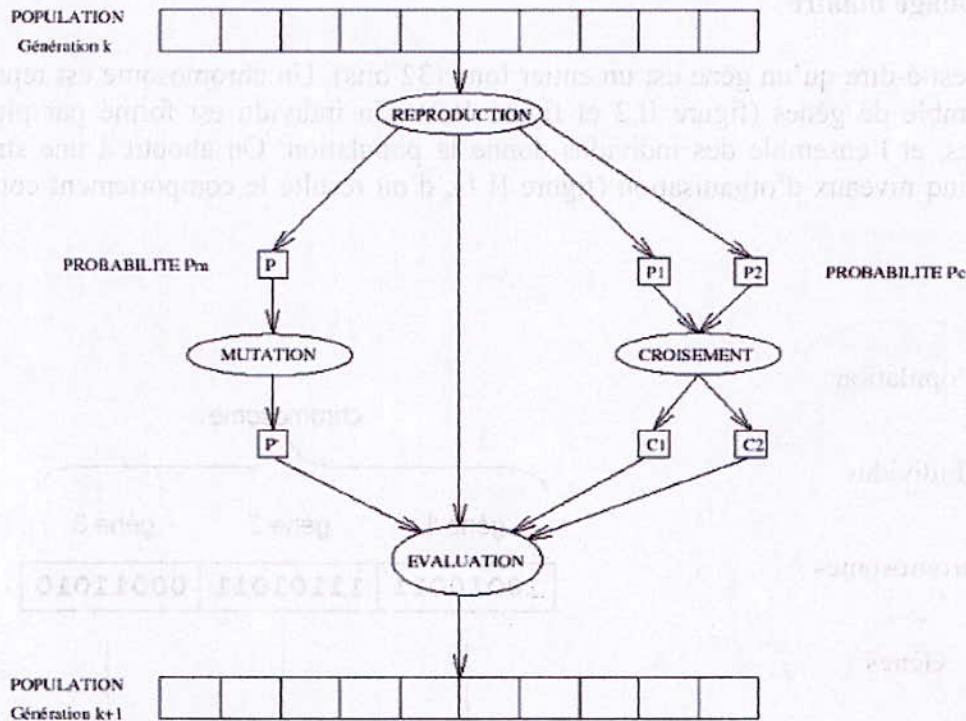


Les algorithmes génétiques (AG) sont des algorithmes itératifs d'optimisation pseudo aléatoire fondés sur les mécanismes de la sélection naturelle et de la génétique. Leur fonctionnement est extrêmement simple, il manipule une population de taille constante. Afin d'illustrer notre raisonnement, nous allons nous placer dans un monde où la génétique est très simplifiée. Les "chromosomes" codent un ensemble de caractéristiques liées. La population est formée de points candidat appelé individu. La taille constante de la population entraîne un phénomène de compétition entre les individus. Chaque individu représente le codage d'une solution potentielle au problème à résoudre, il est constitué d'un ensemble d'éléments appelés chromosome qui, à leur tour, constituent des gènes pouvant prendre plusieurs valeurs appartenant à un alphabet non forcément numérique.

A chaque itération, appelée génération, est créée une nouvelle population avec le même nombre d'individus. Cette génération consiste en des individus mieux « adaptés » à leur environnement tel qu'il est représenté par la fonction sélective. Au fur et à mesure des générations, les individus vont tendre vers l'optimum de la fonction sélective. La création d'une nouvelle population à partir de la précédente se fait par l'application des opérateurs génétiques qui sont : le croisement et la mutation.

Les opérateurs imitent le processus d'hérédité des gènes pour la création des nouveaux descendants à chaque génération. L'opération d'évaluation imite le processus d'évolution Darwinienne pour créer les populations de la génération.





De façon générale, un AG fonctionne de la façon suivante:

**Étape 1:** On génère une population initiale de taille  $n$  chromosomes, puis on choisit au hasard les gènes qui composent chaque chromosome: c'est la première génération de chromosomes.

**Étape 2:** On évalue chaque chromosome par la fonction objectif, ce qui permet de déduire sa valeur d'aptitude.

**Étape 3:** Le cycle de génération des populations commence alors, chaque nouvelle population remplaçant la précédente. Le nombre  $x$  de générations est déterminé au départ. Dans chaque génération, on choisit  $n$  chromosomes auxquels on va appliquer les différents opérateurs génétiques. Après chaque génération, les  $n$  nouveaux chromosomes créés remplacent la génération précédente.

Après la  $x$  ième génération, les chromosomes auront évolué de telle façon que cette dernière génération contienne des chromosomes meilleurs que ceux des générations précédentes.

### II.3. Codage [9] et [10]

Pour résoudre un problème il faut d'abord coder les paramètres, un gène correspond à une variable d'optimisation, et un ensemble de gène correspond à un chromosome un individu a un ou plusieurs chromosomes et une population c'est un ensemble d'individus. Ils existent plusieurs forme de codage nous présenterons deux types de codage ; binaire et réel.

### II.3.1. Codage binaire

C'est-à-dire qu'un gène est un entier long (32 bits). Un chromosome est représenté par un ensemble de gènes (figure II.2 et figure II.3). Un individu est formé par plusieurs chromosomes, et l'ensemble des individus donne la population. On aboutit à une structure présentant cinq niveaux d'organisation (figure II.1), d'où résulte le comportement complexe des AG :

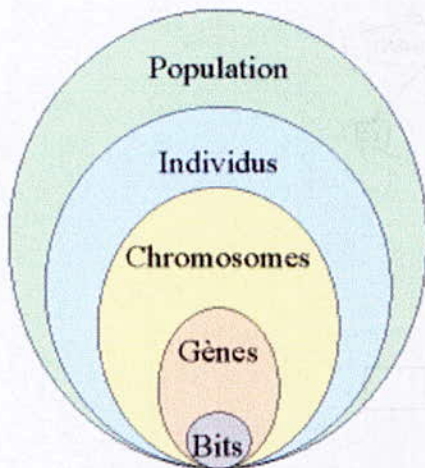


figure II.1 : les cinq niveaux d'organisation de l'algorithme génétique

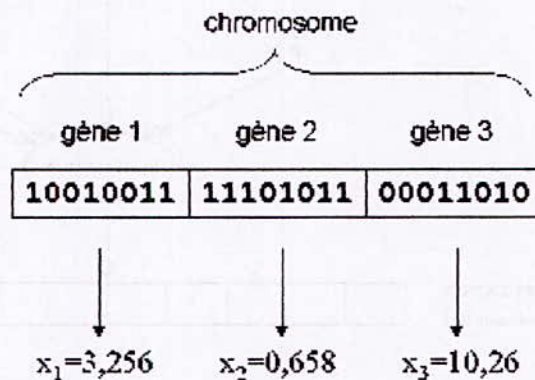


figure II.2: Schéma de codage des variables d'optimisation  $X_i$

Un des avantages du codage binaire est que l'on peut facilement coder toutes sortes d'objets : des réels, des entiers, des valeurs booléennes, des chaînes de caractères... Cela nécessite simplement l'usage de fonctions de codage et décodage pour passer d'une représentation à l'autre.

Supposons que les  $n$  variables sont supposées réelles. Nous considérons un espace de recherche fini :

$$x_{i\min} \leq x_i \leq x_{i\max} \quad \forall i \in [1; n]$$

Afin de coder nos variables réelles en binaire, discrétisons l'espace de recherche. Ainsi un codage sur 32 bits implique une discrétisation des intervalles en  $g_{\max} = 2^{32} - 1 = 4\,294\,967\,295$  valeurs discrètes. Notons au passage que si cette discrétisation est plus fine que celle du modèle physique utilisé, la fonction est assimilable à une fonction escalier, si on la considère à une échelle suffisamment petite.

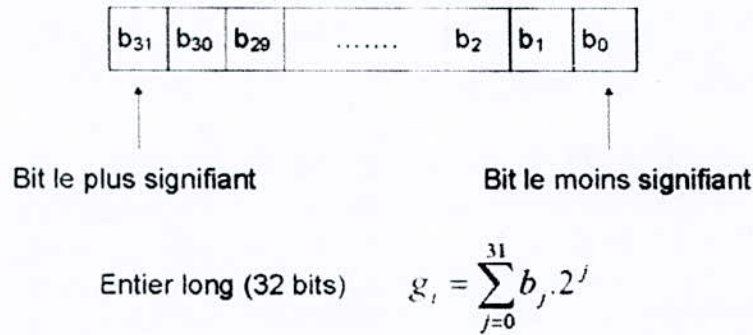


Figure II.3 : codage de gène par un entier long de 32 bits

A chaque variable réelle  $x_i$  on associe donc un entier long  $g_i$  :

$$0 \leq g_i \leq g_{\max} \quad \forall i \in [1, n]$$

Les formules de codage et décodage sont alors les suivantes :

$$g_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \cdot g_{\max}$$

$$x_i = x_{\min} + (x_{\max} - x_{\min}) \cdot \frac{g_i}{g_{\max}}$$

A 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 : 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 six chiffres après la virgule est requise, la taille de l'individu est 3000. Cela, en retour, génère un espace de recherche d'environ  $1000^6$ . Pour de tels problèmes, les algorithmes génétiques basés sur des représentations binaires ont de faibles performances.
- La distance de Hamming (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 la distance de Hamming vaut 3) 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.
- Suivant le problème, la résolution de l'algorithme peut être coûteux en temps.

### II.3.2. Codage réel

Ce codage consiste simplement à la concaténation des variables  $x_i$  d'un individu  $x$ .

Exemple : Un individu  $x$  (25, 31, 8) est codé 25| 31| 8. Ce codage présente des avantages majeurs. Il est plus efficace que le codage binaire avec un espace de recherche identique à l'espace du problème posé. Il a le mérite d'être simple à utiliser, l'évaluation de la formation coût est plus rapide. Le codage réel évite de faire le transcodage du binaire naturel vers les réels à chaque évaluation. Néanmoins, il possède deux inconvénients, son alphabet est infini, et a besoin d'opérateurs appropriés.

### II.4. Fonction de performance (fitness)

Chaque chromosome apporte une solution potentielle au problème à résoudre. Néanmoins, ces solutions n'ont pas toutes le même degré de pertinence. C'est à la fonction de performance (*fitness*) de mesurer cette efficacité pour permettre à l'AG de faire évoluer la population dans un sens bénéfique pour la recherche de la meilleure solution. Autrement dit, la fonction de performance  $f$ , doit pouvoir attribuer à chaque individu un indicateur positif représentant sa pertinence pour le problème qu'on cherche à résoudre.

On peut par exemple admettre la valeur de la fonction objectif ( $f(x) = x^2$ ), celle-ci peut être utilisée directement puisqu'elle est nécessairement positive. De manière plus générale nous représentons la performance d'une séquence  $A$  par une fonction à valeurs positives :

$$f : A \rightarrow f(A) \in \mathbb{R}_+$$

Le but d'un algorithme génétique est alors simplement de trouver la chaîne qui maximise cette fonction  $f$ . Bien évidemment, dans le cas d'un codage binaire, chaque problème particulier nécessitera ses propres fonctions  $f$ .

Une fois que la performance de chaque individu dans la population actuelle est évaluée, les mécanismes évolutionnaires entrent en jeu pour procéder à l'exploration de l'espace des valeurs. Les opérateurs de l'AG cherchent à imiter ces mécanismes.

## II.5. Opérateur de sélection [7] et [9]

### II.5.1. Définition de la sélection

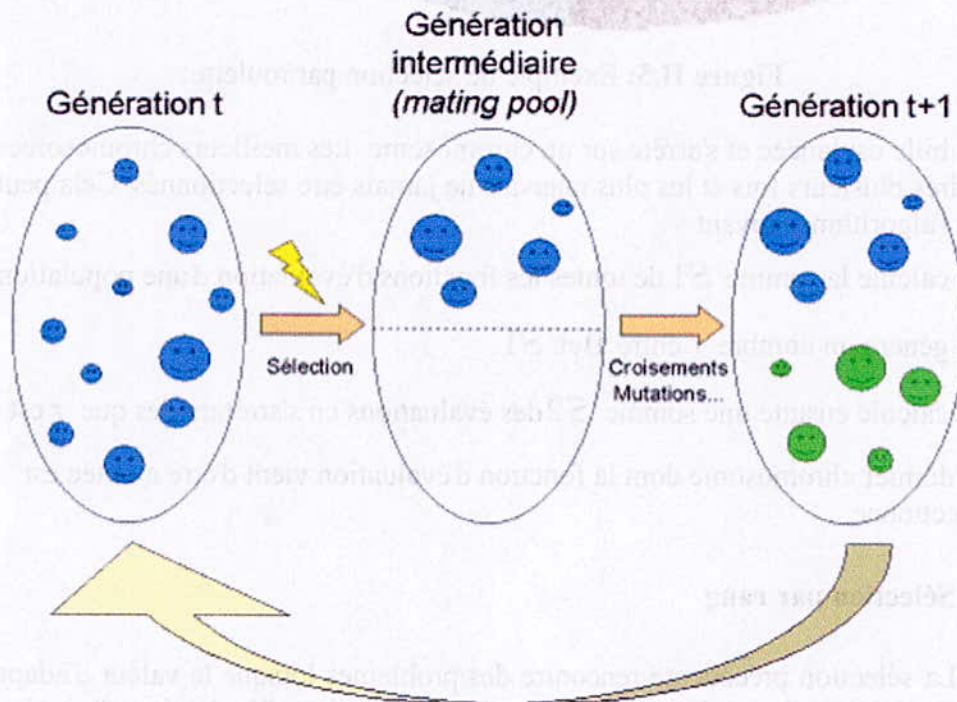
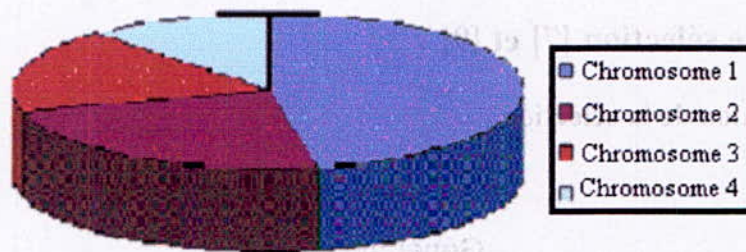


Figure II.4 : représentation schématique du fonctionnement.

Cet opérateur est peut-être le plus important puisqu'il permet aux individus d'une population de survivre, de se reproduire ou de mourir. En règle générale, la probabilité de survie d'un individu sera directement liée à son efficacité relative au sein de la population. Les individus ayant une grande fitness ont donc plus de chance d'être sélectionnés. On parle alors de sélection proportionnelle. La sélection fournit la force motrice dans l'algorithme génétique, elle est basée sur le principe de la sélection darwinienne. La pression de sélection est critique à l'une des extrémités, la recherche de la solution optimale va se terminer prématurément. Tandis qu'à l'autre extrémité, la progression est plus lente que nécessaire. Typiquement, une faible pression est indiquée, au lancement de la recherche, en faveur d'une large exploration de l'espace des solutions, tandis qu'une forte pression de sélection est recommandée, à la fin en vue d'exploiter les régions les plus prometteuses de l'espace des solutions.

### II.5.2. Sélection par roulette (*wheel*)

Les parents sont sélectionnés en fonction de leur performance. Meilleur est le résultat codé par un chromosome, plus grandes sont ses chances d'être sélectionné. Il faut imaginer une sorte de roulette de casino sur laquelle sont placés tous les chromosomes de la population, la place accordée à chacun des chromosomes étant en relation avec sa valeur d'adaptation. Cette roulette est représentée par la figure II.5.



**Figure II.5:** Exemple de sélection par roulette

Ensuite, la bille est lancée et s'arrête sur un chromosome. Les meilleurs chromosomes peuvent ainsi être tirés plusieurs fois et les plus mauvais ne jamais être sélectionnés. Cela peut être simulé par l'algorithme suivant :

1. On calcule la somme  $S1$  de toutes les fonctions d'évaluation d'une population.
2. On génère un nombre  $r$  entre 0 et  $S1$ .
3. On calcule ensuite une somme  $S2$  des évaluations en s'arrêtant dès que  $r$  est dépassé.
4. Le dernier chromosome dont la fonction d'évaluation vient d'être ajoutée est sélectionné.

### II.5.3. Sélection par rang

La sélection précédente rencontre des problèmes lorsque la valeur d'adaptation des chromosomes varie énormément. Si la meilleure fonction d'évaluation d'un chromosome représente 90% de la roulette alors les autres chromosomes auront très peu de chance d'être sélectionnés et on arriverait à une stagnation de l'évolution.

La sélection par rang trie d'abord la population par fitness. Ensuite, chaque chromosome se voit associé un rang en fonction de sa position. Ainsi le plus mauvais chromosome aura le rang 1, le suivant 2, et ainsi de suite jusqu'au meilleur chromosome qui aura le rang  $N$  (pour une population de  $N$  chromosomes). La sélection par rang d'un chromosome est la même que par roulette, mais les proportions sont en relation avec le rang plutôt qu'avec la valeur de l'évaluation. Avec cette méthode de sélection, tous les chromosomes ont une chance d'être sélectionnés. Cependant, elle conduit à une convergence plus lente vers la bonne solution. Ceci est dû au fait que les meilleurs chromosomes ne diffèrent pas énormément des plus mauvais.

### II.5.4. Sélection steady-state

Ce n'est pas une méthode particulière de sélection des chromosomes parents. L'idée principale est qu'une grande partie de la population puisse survivre à la prochaine génération. L'algorithme génétique marche alors de la manière suivante. A chaque génération sont sélectionnés quelques chromosomes (parmi ceux qui ont le meilleur coût) pour créer des chromosomes fils. Ensuite les chromosomes les plus mauvais sont retirés et remplacés par les nouveaux. Le reste de la population survie à la nouvelle génération.

### II.5.5. Sélection par tournoi

Sur une population de  $m$  chromosomes, on forme  $m$  paires de chromosomes. Dans les paramètres de l'AG, on détermine une probabilité de victoire du plus fort. Cette probabilité représente la chance qu'a le meilleur chromosome de chaque paire d'être sélectionné. Cette probabilité doit être grande (entre 70% et 100%). A partir des  $m$  paires, on détermine ainsi  $m$  individus pour la reproduction.

### II.6. Opérateur de croisement [8] et [11]

L'opérateur de croisement permet la création de nouveaux individus selon un processus très simple. Il permet donc l'échange d'information entre les individus.

Tout d'abord, deux individus, qui forment alors un couple, sont tirés au sein de la nouvelle population issue de la reproduction. Puis un (potentiellement plusieurs) site de croisement est tiré aléatoirement (chiffre entre 0 et 1). Enfin, selon une probabilité de croisement  $p_c$  qui est définie comme le pourcentage de nombre espéré des individus qui vont subir l'opération de croisement, par rapport à la taille de la population. Une probabilité de croisement élevée permet une exploration plus large de l'espace de solutions et réduit les chances de fixation sur un faux optimum. Mais si cette probabilité est trop élevée, alors une quantité importante de temps de calcul risque d'être perdu dans l'exploration des régions non prometteuses de l'espace de solutions. Typiquement la probabilité de croisement varie entre 0.6 et 1.0.

Cet opérateur permet la création de deux nouveaux individus. Toutefois, un individu sélectionné lors de la reproduction ne subit pas nécessairement l'action d'un croisement. Ce dernier ne s'effectue qu'avec une certaine probabilité. Plus cette probabilité est élevée et plus la population subira de changement.

#### II.6.1. Croisement binaire

##### II.6.1.1. Croisement en un point

Une position de  $k$  est choisie uniformément entre 1 et  $(l-1)$ , deux nouvelles chaînes sont créées en échangeant tous les caractères entre les positions  $(k+1)$  et  $l$  de chaque chaîne de la paire considérée, les nouvelles chaînes peuvent être totalement différentes de leurs parents. A remarquer qu'un croisement n'aura aucune conséquence sur un bit dont les parents ont la même position.



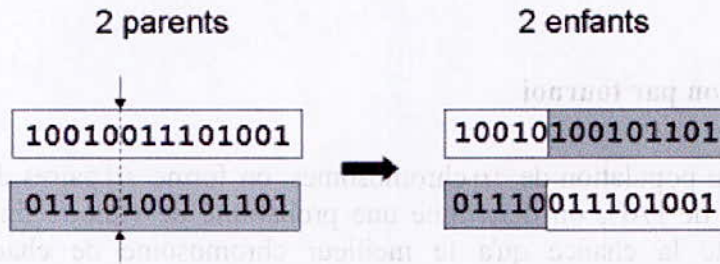


Figure II.6 : représentation schématique du croisement en 1 point

**II.6.1.2. Croisement en deux points**

Le même principe du croisement en un point il est appliqué au niveau du croisement en deux points à la différence que deux coupures sont choisies au hasard comme il est schématisé dans la figure II.7. Cette opération est généralement considérée comme plus efficace que la précédente.

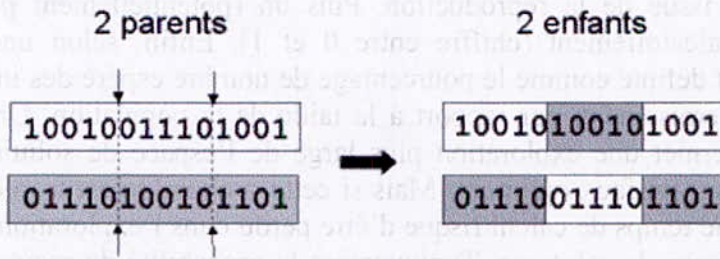


Figure II.7 : représentation schématique du croisement en 2 points.

**II.6.1.3. Croisement uniforme**

Chaque bit du descendant est créé à partir du bit correspondant de l'un des parents. Les bits sont choisis selon un vecteur appelé vecteur masque de croisement. Lorsqu'il y a un 1 dans le vecteur masque, le bit est copié à partir du premier parent est lorsqu'il y a un 0 le bit est copié à partir du second parent comme il est montré sur la fig II.8.

Le procédé est répété avec d'autres parents pour produire le second descendant. Un nouveau vecteur masque est généré aléatoirement pour chaque paire des parents.

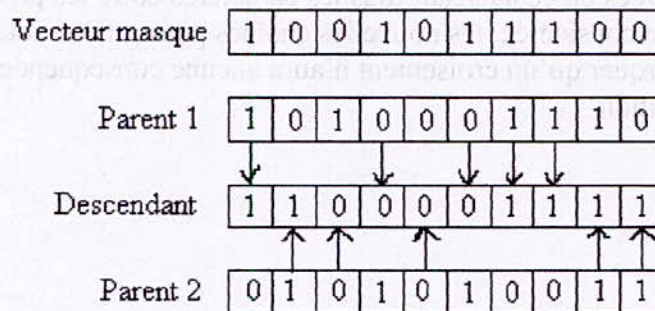


Figure II.8 : représentation schématique du croisement uniforme



## II.6.2. Croisement réel

### II.6.2.1. Croisement simple

Cette méthode est identique dans le principe à celle décrite auparavant. Pour ce faire, nous générons un nombre aléatoire  $r$  à partir d'une distribution uniforme sur l'ensemble de  $f$ , et deux nouveaux individus,  $X'$  et  $Y'$  sont créés selon la règle suivante :

$$X' = \begin{cases} X_i & \text{Si } i < r \\ Y_i & \text{Sinon} \end{cases}$$

$$Y'_i = \begin{cases} Y_i & \text{Si } i < r \\ X_i & \text{Sinon} \end{cases}$$

### II.6.2.2. Croisement arithmétique

Ce croisement effectue une simple combinaison linéaire entre les parents. Soit, après avoir généré un chiffre aléatoire,  $\alpha = U(0,1)$ , les nouveaux parents sont :

$$\begin{aligned} X' &= \alpha X + (1-\alpha)Y \\ Y' &= (1-\alpha)X + \alpha Y \end{aligned}$$

### II.6.2.3. Croisement heuristique

Enfin, il existe aussi le croisement heuristique. Cet opérateur effectue une extrapolation linéaire des individus. Un nouvel individu,  $X'$  est créé selon le processus suivant (sous l'hypothèse que  $X > Y$  en terme de fitness, sinon nous inversons  $X$  et  $Y$  dans les équations) :

$$\begin{aligned} X' &= X + r(X - Y) \\ Y' &= X \end{aligned}$$

Ou/et

$$\text{faisabilité} = \begin{cases} 1 & \text{Si } b_1^i < X_i < b_2^i \quad \forall i \\ 0 & \text{Sinon} \end{cases}$$

Ou  $b_1^i$  et  $b_2^i$  sont des bornes autorisées pour  $X_i$ , et avec  $r$  un nombre aléatoire tiré dans  $U(0;1)$ . Nous devons donc avoir tout le temps  $X_i \in [b_1^i; b_2^i]$ . Si  $X'$  n'est pas faisable

(faisabilité nulle) alors un nombre  $r$  est relié et la procédure est recommencée jusqu'à ce que la solution soit faisable ou qu'un certain nombre d'essais ait été effectué.

Le cas où  $f(X)=f(Y)$  (même fitness) on reproduit simplement  $X$  et  $Y$ . cet opérateur est le seul des croisements à utiliser directement une information reliée à la fitness. Cet opérateur est un croisement unique pour les raisons suivantes :

- Il utilise les valeurs de la fonction objective afin de déterminer une direction de recherche.
- Il produit seulement un enfant et il peut ne produire aucun enfant.

Il semble que le croisement heuristique contribue à trouver une solution plus précise : ses principales responsabilités dans la recherche de la solution c'est de faire une recherche dans une direction prometteuse.

#### II.6.2.4. Croisement MPX

Ce croisement conserve dans le fils  $i$  la zone interne de parent  $i$  (zone comprise entre  $N1$  et  $N2$  tirés aléatoirement). Ensuite, pour les cases vides restantes, le  $k$  ième chromosome de fils  $i$  sera rempli avec le  $k$  ième de parent  $j$  s'il ne provoque pas de doublon. A défaut, on essaiera le  $k$  ième chromosome de parent  $i$  ou enfin le premier chromosome de la zone interne de parent  $j$  qui ne provoque pas de doublon. Ce croisement est illustré sur la figure.II.9.

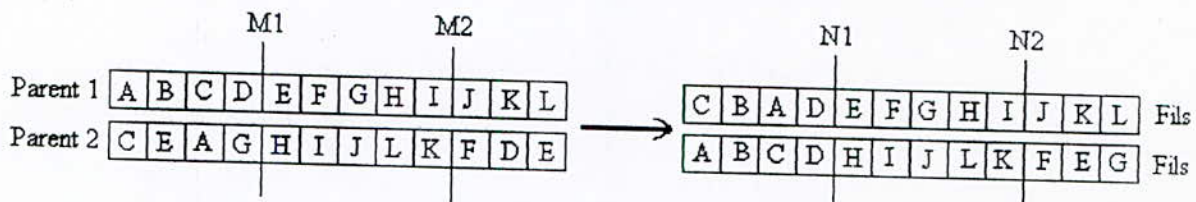


Figure.II.9 : Croisement MPX

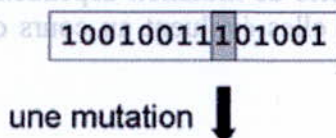
## II.7. Opérateur de mutation [11] et [12]

La mutation est le processus selon lequel la valeur d'un gène choisi au hasard dans un chromosome est régénérée. Ce processus ne survient qu'occasionnellement dans un algorithme génétique. En modifiant aléatoirement la valeur d'un bit dans une chaîne, ou une variable réelle qui est de nouveau tirée au hasard sur son intervalle. La mutation est utile pour ramener du matériel génétique qui aurait été oublié par les opérateurs de sélection et de croisement.

### II.7.1. Mutation binaire

Nous définissons une mutation binaire comme étant l'inversion d'un bit dans un chromosome (Fig.II.10 ). Cela revient à modifier aléatoirement la valeur d'un paramètre du dispositif. Les mutations jouent le rôle de bruit et empêchent l'évolution de se figer. Elles permettent d'assurer une recherche aussi bien globale que locale, selon le poids et le nombre

des bits mutés. De plus, elles garantissent mathématiquement que l'optimum global peut être atteint.



10010011001001

Fig. II.10 : représentation schématique d'une mutation dans un chromosome.

D'autre part, une population trop petite peut s'homogénéiser à cause des erreurs stochastiques : les gènes favorisés par le hasard peuvent se répandre au détriment des autres. Cet autre mécanisme de l'évolution, qui existe même en l'absence de sélection, est connu sous le nom de dérive génétique. Du point de vue du dispositif, cela signifie que l'on risque alors d'aboutir à des dispositifs qui ne seront pas forcément optimaux. Les mutations permettent de contrebalancer cet effet en introduisant constamment de nouveaux gènes dans la population.

De nombreuses méthodes existent. Souvent la probabilité de mutation  $P_m$  par bit et par génération est fixée entre 0,001 et 0,01. On peut prendre également  $p_m = 1/l$  où  $l$  est le nombre de bits composant un chromosome. Il est possible d'associer une probabilité différente de chaque gène. Et ces probabilités peuvent être fixes ou évoluer dans le temps.

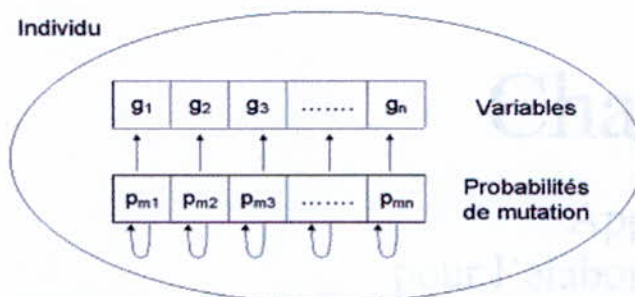


Fig. II.11 : principe de l'auto-adaptation. A chaque variable est associée sa propre probabilité de mutation, qui est elle-même soumise au processus d'évolution. L'individu possède donc un second chromosome codant ces probabilités.

Après divers essais, ont abouti à la méthode d'*auto-adaptation* des probabilités de mutation. Si dans un environnement stable il est préférable d'avoir un taux de mutation faible, la survie d'une espèce dans un environnement subissant une évolution rapide nécessite un taux de mutation élevé permettant une adaptation rapide. Les taux de mutation d'une espèce dépendent donc de leur environnement.

Pour prendre en compte cette formulation biologique et l'adapter à notre cas, ils ont introduit dans chaque individu (dispositif) un second chromosome (ensemble de paramètres) dont les gènes (paramètres) représentent les probabilités de mutation de chaque gène du premier chromosome (Fig. II.11). Ce second chromosome est géré de façon identique au premier, c'est-à-dire qu'il est lui-même soumis aux opérateurs génétiques (croisement et mutation). Cela revient à fixer les probabilités assurant la modification des valeurs des paramètres du composant en fonction des valeurs d'un ensemble d'autres paramètres (les probabilités de mutation). Au cours du déroulement de l'algorithme, les gènes et les individus ayant des probabilités de mutation trop élevées ont tendance à disparaître. De même, les gènes

### III.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.

Concernant notre études, nous avons choisi un modèle assez simple qui est le même pour les trois propriétés étudiées.

$$p = a T_{eb} + b 10^3 d_{20} + c$$

$p$  : une des trois propriétés critiques étudié ( $T_c(K), P_c(\text{bar}), V_c(\text{cm}^3/\text{mol})$ ).

$a, b$  et  $c$  : les coefficients du modèle.

$d_{20}$  : densité du corps purs à 20°C

$T_{eb}$  : température d'ébullition du corps purs en kelvin (K).

L'espace de recherche des coefficients pour ce modèle est défini dans le tableau ci-dessous :

$T_c$	$0 < a < 1$	$0 < b < 1$	$0 < c < 100$
$P_c$	$-0.1 < a < 0$	$0 < b < 0.1$	$0 < c < 100$
$V_c$	$-5 < a < 5$	$-5 < b < 5$	$0 < c < 300$

Ce choix a été guidé par une estimation des valeurs des coefficients ainsi que par plusieurs tâtonnements des autres espaces de recherches possibles.

La structure du programme est constituée d'un programme principal qui fait appel à plusieurs fonctions ainsi que d'une boucle qui s'arrête une fois les conditions satisfaites, aussi dans notre programme nous avons trois entrées (taille de la population, l'erreur maximale souhaitée et la propriété désirée)

Pour une telle réalisation, nous avons pris une banque de données de trois types de familles (paraffine, naphène et aromatique) sous forme de vecteurs (voir Annexe.I).

Exemple pour la température critique :

Un vecteur pour les valeurs de la température d'ébullition  $T_{eb}$

Un vecteur pour les valeurs de la densité  $d$

Un vecteur pour les valeurs de la température critique  $T_c$

### III.2. Codage réel des paramètres d'entrées

Pour résoudre notre problème nous avons choisi de coder les valeurs des coefficients de notre modèle (de chaque famille 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. Ce choix a été motivé par la simplicité du codage (vue qu'il n'y a pas de transformation en binaire), et de la programmation des opérateurs (sélection, croisement, mutation).

### III.3. Création de la population initiale

Dans cette étape nous choisissons en premier lieu la taille de 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 POP1, le nombre de colonne de cette matrice est de 3 (qui représente le nombre de coefficient). Ainsi nous allons nous retrouver avec une matrice de taille  $n*m$  tel que  $n$  : est la taille de la population et  $m$  : nombre de coefficients du modèle choisi (pour notre exemple  $m=3$ ), 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 une fonction (la *fonction genererpop*) qui va générer ces valeurs, par exemple pour  $T_c$ , l'algorithme est le suivant :

Pour  $i=1$  à  $x$

$$\text{Pop1}(i, 1) = \text{Rand}$$

$$\text{Pop1}(i, 2) = \text{Rand}$$

$$\text{Pop1}(i, 3) = 100 * \text{Rand}$$

$x$  : est la taille de la population.

Rand : une valeur tirée au hasard (Rand est définie dans la banque de données du logiciel Matlab) comprise entre 0 et 1.

Donc le programme va générer des éléments de la matrice POP1 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 enfin la troisième colonne représente le coefficient  $c$ . Nous allons donc nous retrouver avec  $x$  corrélations ( $x$  étant la taille de la population) pour chaque propriété. Cette matrice générée est appelée population initiale.

#### Exemple de création d'une population :

Soit à optimiser les coefficients du modèle suivant :

$$Z = a X + b Y + c$$

Avec une banque de données suivante :

$$\begin{aligned} X &= [2 \ 4 \ 5 \ 6] \\ Y &= [0.2 \ 0.3 \ 0.5 \ 0.8] \\ Z &= [12 \ 15 \ 17 \ 20] \end{aligned}$$

Nous aurons donc à générer une population de quatre individus à trois coefficients, avec un espace de recherche définie comme suit pour les trois coefficients

$2 > a > 0$ ,  $6 > b > 4$  et  $9 > c > 7$ . Bien sûr le coefficient (a) représente la colonne numéro un dans notre matrice POP1 et (b) la deuxième colonne, et (c) la troisième colonne.

Pour  $i=1$  à 4

$$\begin{aligned} \text{POP1}(i, 1) &= 2 * \text{Rand} \\ \text{POP1}(i, 2) &= 2 * \text{Rand} + 4 \\ \text{POP1}(i, 3) &= 2 * \text{Rand} + 7 \end{aligned}$$

Nous aurons la matrice suivante :

$$\text{POP1} = \begin{pmatrix} 0.11 & 5.2 & 7.1 \\ 0.83 & 4.61 & 8.74 \\ 0.03 & 5.53 & 8.94 \\ 1.98 & 5.57 & 7.87 \end{pmatrix}$$

### III.4. 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 une fonction qui est appelée à partir du programme principal, *fonction fitness*. Le rôle de cette fonction est de calculer pour chaque ligne la propriété critique à partir du modèle et ceci pour chaque  $T_{eb}$  et  $d$  de la banque de donnée, et de la comparer avec celle du modèle.

$$E_i = 100 * \text{valeurs absolue} (T_{ci}(\text{modèle}) - T_{ci}(\text{banque de donnée})) / T_{ci}(\text{banque de donnée})$$

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  $\text{Max}(E)$ . 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 4 valeurs de  $T_c$  correspondante à 4 valeurs de  $T_{eb}$  et de  $d$  nous prendrons la valeurs du fitness qui correspond au maximum des valeurs  $E$ , donc la valeur du fitness de cette ligne sera égale à :

$$\text{Fitness}(i) = \max(E_1, E_2, E_3, E_4)$$

$i$  : 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.

Pour l'exemple cité précédemment, Essayons de calculer les valeurs du fitness de la matrice POP1.

$$POP1 = \begin{pmatrix} 0.11 & 5.2 & 7.1 \\ 0.83 & 4.61 & 8.74 \\ 0.03 & 5.53 & 8.94 \\ 1.98 & 5.57 & 7.87 \end{pmatrix}$$

Calcul des erreurs  $E_i$  et du fitness pour chaque ligne

*Pour la première ligne :*

$$E = (30.1799 \quad 39.0952 \quad 39.4398 \quad 40.1234)$$

$$\text{Max}(E) = 40.1234$$

*Pour la deuxième ligne :*

$$E = (5.5647 \quad 10.3018 \quad 10.5442 \quad 12.8938)$$

$$\text{Max}(E) = 12.8938$$

*Pour la troisième ligne*

$$E = (15.7591 \quad 28.5164 \quad 30.2369 \quad 32.2474)$$

$$\text{Max}(E) = 32.2474$$

*Pour la quatrième ligne :*

$$E = (7.9433 \quad 16.4753 \quad 20.9824 \quad 21.1024)$$

$$\text{Max}(E) = 21.1024$$

*Nous obtiendrons ainsi le vecteur du fitness suivant :*

$$\text{Fitness} = (40.1234 \quad 12.8938 \quad 32.2474 \quad 21.1024)$$

### III.5. Opérateur de sélection

C'est le premier opérateur exécuté par le programme principal, ainsi il fait appel à une fonction (*fonction select*) sélectionnant les meilleurs individus pour les deux autres opérateurs qui vont suivre. Dans ce sous programme on calcul en premier lieu deux probabilités  $p$  et  $q$  que nous définissons comme suit :

$$P(i) = \frac{\text{Fitness}(i)}{\sum_{j=1}^x \text{Fitness}(j)} \qquad q(i) = \sum_{j=1}^i P(j)$$

$i$  : 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 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 supérieure et copie la ligne de la POP1 qui correspond à cet indice (c'est la ligne choisi), sur la ligne qui correspond à l'indice de l'élément du vecteur  $r_s$  border (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 *select* est :

```

Pour i=1 à x
  Pour j=1 à x
    Si  $r_s(i) < q(j)$  alors
      Si  $\text{fitness}(j) < \text{fitness}(i)$  alors
        Pour k=1 à 3
           $\text{POP1}(i,k) = \text{POP1}(j,k)$ 
        Fin
      Sinon choisir la borne suivante
    Fin
  Fin
Fin

```

Nous obtiendrons une nouvelle matrice POP'1.



Toujours pour le même exemple et pour la population de quatre individus :

$$POP_1 = \begin{pmatrix} 0.11 & 5.2 & 7.1 \\ 0.83 & 4.61 & 8.74 \\ 0.03 & 5.53 & 8.94 \\ 1.98 & 5.57 & 7.87 \end{pmatrix} \quad \text{Fitness} = \begin{pmatrix} 40.1234 \\ 12.8938 \\ 32.2474 \\ 21.1024 \end{pmatrix}$$

Somme des fitness = 82.2646 et probabilité  $p$  est donc :

$$p(1) = 40.1234 / 82.2646 = 0.3772$$

$$p(2) = 12.8938 / 82.2646 = 0.1212$$

$$p(3) = 32.2474 / 82.2646 = 0.3032$$

$$p(4) = 21.1024 / 82.2646 = 0.1984$$

La probabilité  $q$

$$q(1) = p(1) = 0.3772$$

$$q(2) = p(1) + q(2) = 0.4984$$

$$q(3) = p(1) + q(2) + q(3) = 0.8016$$

$$q(4) = p(1) + q(2) + q(3) + q(4) = 1$$

$$POP_1 = \begin{pmatrix} 0.11 & 5.2 & 7.1 \\ 0.83 & 4.61 & 8.74 \\ 0.03 & 5.53 & 8.94 \\ 1.98 & 5.57 & 7.87 \end{pmatrix} \quad \text{Fitness} = \begin{pmatrix} 40.1234 \\ 12.8938 \\ 32.2474 \\ 21.1024 \end{pmatrix} \quad q_s = \begin{pmatrix} 0.3772 \\ 0.4984 \\ 0.8016 \\ 1.0000 \end{pmatrix} \quad r_s = \begin{pmatrix} 0.4983 \\ 0.2140 \\ 0.6435 \\ 0.3200 \end{pmatrix}$$

$q_s(1) < r_s(1) < q_s(2)$  donc il aura une modification de ligne, la ligne 1 deviendra la même que la ligne 2.

$r_s(2) < q_s(1)$  il n'y aura pas de changement puisque  $\text{Fitness}(2) < \text{Fitness}(1)$ , ainsi de suite pour les autres ligne (pour  $q_s(2)$ ,  $q_s(3)$ ,  $q_s(4)$ ).

$q_s(2) < r_s(3) < q_s(3)$  il n'y aura pas de changement, puisque c'est la même ligne.

$r_s(3) < q_s(4)$  il y aura une modification, vue que  $\text{fitness}(4) < \text{fitness}(3)$ , la ligne 3 deviendra la ligne 4.

$r_s(4) < q_s(1)$  il n'y aura pas de changement de ligne puisque  $\text{Fitness}(1) > \text{Fitness}(4)$ .

$r_s(4) < q_s(2)$  nous aurons un changement ( $\text{fitness}(4) < \text{fitness}(3)$ ), la ligne 4 prendra les mêmes valeurs que la ligne 2.

Notre matrice deviendra donc comme suit (le programme va recalculer les fitness de chaque ligne).

$$POPI' = \begin{pmatrix} 0.83 & 4.61 & 8.74 \\ 0.83 & 4.61 & 8.74 \\ 1.98 & 5.57 & 7.87 \\ 0.83 & 4.61 & 8.74 \end{pmatrix} \quad \text{Fitness} = \begin{pmatrix} 12.8938 \\ 12.8938 \\ 21.1024 \\ 12.8938 \end{pmatrix}$$

Nous avons donc éliminé les deux mauvaises valeurs du fitness c'est-à-dire les corrélations aux coefficients donnant les plus grandes erreurs.

### III.6. 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 une fonction (fonction *Cross*) qui va croiser soit en un point ou en deux points, ou plus (tous dépendra de la probabilité  $p_c$ ), ou du nombre de coefficient 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 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 pour d'autres types de corrélations nous serions 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  $\text{rand}$  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 (deux individus pour une opération de croisement). 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 la probabilité  $p_c$ . Pour chaque couple d'individus, un autre vecteur  $r'_c$  est créé composé de trois éléments  $\text{rand}$  (trois 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 (couple) est permutée, ainsi de suite pour les deux autres colonnes.

Les cas possibles pour notre corrélation sont les suivants :

*Permutation entre la première colonne*

$$\begin{array}{|c|c|c|} \hline a & b & c \\ \hline a' & b' & c' \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|c|} \hline a' & b & c \\ \hline a & b' & c' \\ \hline \end{array}$$

*Permutation entre la deuxième colonne*

$$\begin{array}{|c|c|c|} \hline a & b & c \\ \hline a' & b' & c' \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|c|} \hline a & b' & c \\ \hline a' & b & c' \\ \hline \end{array}$$

*Permutation entre la troisième colonne*

$$\begin{array}{|c|c|c|} \hline a & b & c \\ \hline a' & b' & c' \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|c|} \hline a & b & c' \\ \hline a' & b' & c \\ \hline \end{array}$$

*Permutation entre la première et la deuxième colonne*

$$\begin{array}{|c|c|c|} \hline a & b & c \\ \hline a' & b' & c' \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|c|} \hline a' & b' & c \\ \hline a & b & c' \\ \hline \end{array}$$

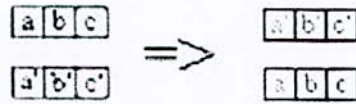
*Permutation entre la deuxième et la troisième colonne*

$$\begin{array}{|c|c|c|} \hline a & b & c \\ \hline a' & b' & c' \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|c|} \hline a & b' & c' \\ \hline a' & b & c \\ \hline \end{array}$$

*Permutation entre la première et la troisième colonne*

$$\begin{array}{|c|c|c|} \hline a & b & c \\ \hline a' & b' & c' \\ \hline \end{array} \Rightarrow \begin{array}{|c|c|c|} \hline a' & b & c' \\ \hline a & b' & c \\ \hline \end{array}$$

Croisement au niveau des trois colonnes en même temps : ce cas de permutation, est rare au vue d'avoir trois valeurs de rand < 0.75



Revenons à notre exemple ou nous allons subir à la matrice POP 1' une opération de croisement.

$$POP1' = \begin{pmatrix} 0.83 & 4.61 & 8.74 \\ 0.83 & 4.61 & 8.74 \\ 1.98 & 5.57 & 7.87 \\ 0.83 & 4.61 & 8.74 \end{pmatrix} \quad \text{Fitness} = \begin{pmatrix} 12.8938 \\ 12.8938 \\ 21.1024 \\ 12.8938 \end{pmatrix}$$

Le programme va générer le vecteur  $r_c$  d'éléments rand.

$$r_c = \begin{pmatrix} 0.9601 \\ 0.7266 \\ 0.4120 \\ 0.7446 \end{pmatrix}$$

Pour le croisement, le programme va choisir les trois lignes 2, 3 et 4 de la matrice POP1' correspondant respectivement à  $r_c(2)$ ,  $r_c(3)$  et  $r_c(4)$  vue que tous les trois ont une valeurs inférieures à pc ( est qui égale à 0.75 ).

Etant donnée que ce croisement s'effectue deux par deux, dans le cas de cet exemple donc ; le programme laissera tomber la quatrième ligne, en fin de compte nous aurons un croisement uniquement au niveau de la ligne 2 et 3.

Ensuite le programme va crée un vecteur  $r_c'$ .

$$r_c' = (0.2679 \quad 0.4399 \quad 0.9334)$$

Ce qui veut dire que les colonnes choisies pour être permutés sont la colonne une et deux, c'est un croisement en un point. Notre matrice devient comme suit :

$$POP1'' = \begin{pmatrix} 0.83 & 4.61 & 8.74 \\ 1.98 & 5.57 & 8.74 \\ 0.83 & 4.61 & 7.87 \\ 0.83 & 4.61 & 8.74 \end{pmatrix} \quad \text{Fitness} = \begin{pmatrix} 12.8938 \\ 26.1084 \\ 17.2509 \\ 12.8938 \end{pmatrix}$$

Nous obtiendrons ainsi une nouvelle matrice  $POP1''$ , et le fitness obtenu est bien sûr un nouveau fitness calculé. Nous remarquons que nous avons obtenu deux valeurs différentes du fitness, même si elles plus sont grandes que les précédentes mais le but est toujours de trouver une combinaison qui fait évoluer notre génération.

### III.7. Opérateur de Mutation

C'est le dernier opérateur que la programme principal exécute, et c'est probablement le plus important puisque c'est lui qui fait évoluer la génération. La mutation s'exécute par une fonction que l'on appellera *fonction mut*, et qui choisit un des éléments de la matrice  $POP1''$  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'éléments rand tel que la taille de ce vecteur est égale à : taille de la population \* nombre de coefficient. Ce vecteur est important vue que c'est lui qui va choisir les éléments qui vont être muter, et cela en testons chaque élément de ce vecteur, si l'élément est inférieur à notre probabilité de mutation  $p_m$ ; alors l'élément de notre matrice  $POP1''$  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  $POP1''$  qui correspond à l'indice de l'élément de ce vecteur le programme exécute un calcul de la ligne et de la colonne d'après cette formule.

$$j = \text{mod}(k, 3) \\ i = 1 + ((k-j)/3)$$

$j$  : l'indice de la colonne de la matrice  $POP1''$  ( ne peut être > 3)

$i$  : l'indice de la ligne de la matrice  $POP1''$

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

$\text{mod}$  : c'est une fonction qui existe dans la banque de donnée de Matlab et donne le reste de la division de  $k/3$ ; par exemple  $\text{mod}(5,3) = 2$ .

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 de la *fonction mute* pour  $T_c$  par exemple :

Pour  $k=1$  à  $x*3$

Si  $r_m < 0.33$  alors

$j = \text{mod}(k,3)$

Si  $j=0$  alors

$j=3$

Fin

$i = 1 + ((k-j)/3)$

Si  $j > 1$  alors

Si  $j > 2$  alors

$\text{POP1}'(i,j) = 10 * \text{Rand} + 20$

Sinon

$\text{POP1}'(i,j) = 10 * \text{Rand} + 10$

Fin

Sinon

$\text{POP1}'(i,j) = 10 * \text{Rand}$

Fin

Fin

Fin

Pour notre exemple

Soit la matrice  $\text{POP1}''$  qui va subir une opération de mutation tel que ( $p_m = 0.33$ )

$$\text{POP1}'' = \begin{pmatrix} 0.83 & 4.61 & 8.74 \\ 1.98 & 5.57 & 8.74 \\ 0.83 & 4.61 & 7.87 \\ 0.83 & 4.61 & 8.74 \end{pmatrix} \quad \text{Fitness} = \begin{pmatrix} 12.8938 \\ 26.1084 \\ 17.2509 \\ 12.8938 \end{pmatrix}$$

Création du vecteur  $r_m$  tel que sa taille est égale à  $4 * 3 = 12$

$$r_m = \begin{pmatrix} 0.9501 \\ 0.2311 \\ 0.6068 \\ 0.4860 \\ 0.8913 \\ 0.7621 \\ 0.4565 \\ 0.0185 \\ 0.8214 \\ 0.4447 \\ 0.6154 \\ 0.7919 \end{pmatrix}$$

Nous remarquons que deux éléments seulement de ce vecteur sont inférieurs à 0.33 :  $r_m(2)$  et  $r_m(8)$ , le programme va donc calculer les lignes et les colonnes qui correspondes à ces deux éléments du vecteur.

Pour  $r_m(2)$

$$j = \text{mod}(2,3) = 2$$

$$i = 1 + (2-2)/3 = 1$$

donc l'élément choisi pour une mutation est POP1''(1,2).

Pour  $r_m(8)$

$$j = \text{mod}(8,3) = 2$$

$$i = 1 + (8-2)/3 = 3$$

donc l'élément choisi pour une mutation est POP1''(3,2).

Pour POP1''(1,2)

Le programme va tirer une valeur du domaine de recherche comme suite  
 $\text{POP}''(1,2) = 2 * \text{rand} + 4 = 5.8436$

Pour POP1''(3,2)

Le programme va tirer une valeur du domaine de recherche comme suite  
 $\text{POP}''(3,2) = 2 * \text{rand} + 4 = 5.7442$

Nous obtiendrons la matrice suivante :

$$POP1''' = \begin{pmatrix} 0.83 & 5.84 & 8.74 \\ 1.98 & 5.57 & 8.74 \\ 0.83 & 5.47 & 7.87 \\ 0.83 & 4.61 & 8.74 \end{pmatrix} \quad \text{Fitness} = \begin{pmatrix} 7.9610 \\ 26.1084 \\ 14.3798 \\ 12.8938 \end{pmatrix}$$

### III.8. Test d'arrêt

Après l'exécution de tous ces opérateurs nous obtiendrons une nouvelle génération d'individus. Le teste d'arrêt est une simple vérification du fitness de chaque ligne de notre matrice, notre programme repaire 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 une autre fonction qui fait seulement appliquer le modèle ainsi obtenu. Sinon le programme refait tous les calculs avec les différents opérateurs mais cette fois avec une matrice initiale POP1'''. 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) seulement pour une tel exécution il faudra ajouter ces deux paramètres d'entré (voir III.1).

L'algorithme est le suivant :

Trouver le minimum de fitness

Si minimum < E alors

Exécuter le sous programme de l'application

Sinon

ré-exécuter les opérateur avec la matrice initiale POP1'''

Fin

Tel que E est l'erreur choisie.



# Chapitre IV

Résultats et commentaires

## IV.1. Introduction

L'exécution du programme principal s'effectue dans le workspace du Matlab voir la Figure.IV.1, le workspace est une fenêtre ou on fait appelle a des programmes pour l'exécution, aussi c'est la fenêtre ou les résultats s'affichent.

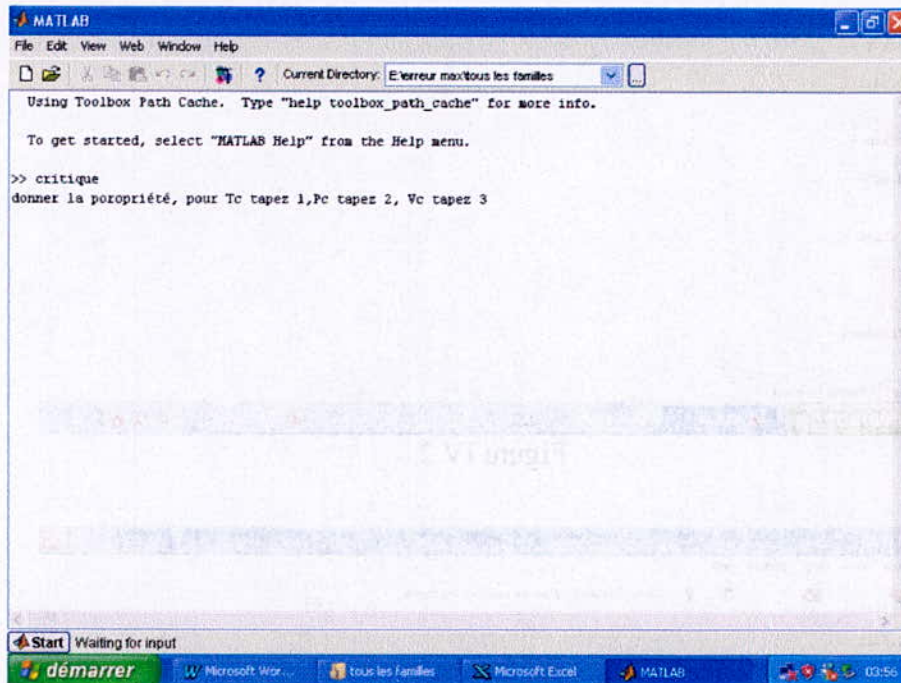


Figure IV.1

Pour l'exécution de notre programme, on fait appel à notre fonction principale on tapant *critique*, lors de l'exécution nous devant choisir la propriété comme il est démontré sur la FigureIV.1, on tape 1 pour  $T_c$ , 2 pour  $P_c$ , et enfin 3 pour  $V_c$ .

L'exécution vise à minimiser l'erreur maximale pour les trois familles en même temps c'est-à-dire une corrélation pour les trois familles. Pour atteindre notre but il nous a fallut une population initiale de 1000 individus (voir III.3).

## IV.2. Résultats obtenus

### IV.2.1. Température critique

Nous avons fixé pour la température critique une erreur maximale de 1.75 % et cela après plusieurs essais; le programme converge après 221 générations et un temps d'exécution inférieur à cinq minutes (Figure IV.2), l'erreur maximale obtenue est de 1.7484 %, l'erreur moyenne est de 0.5412 % et enfin l'erreur minimale est de 0.0036 %

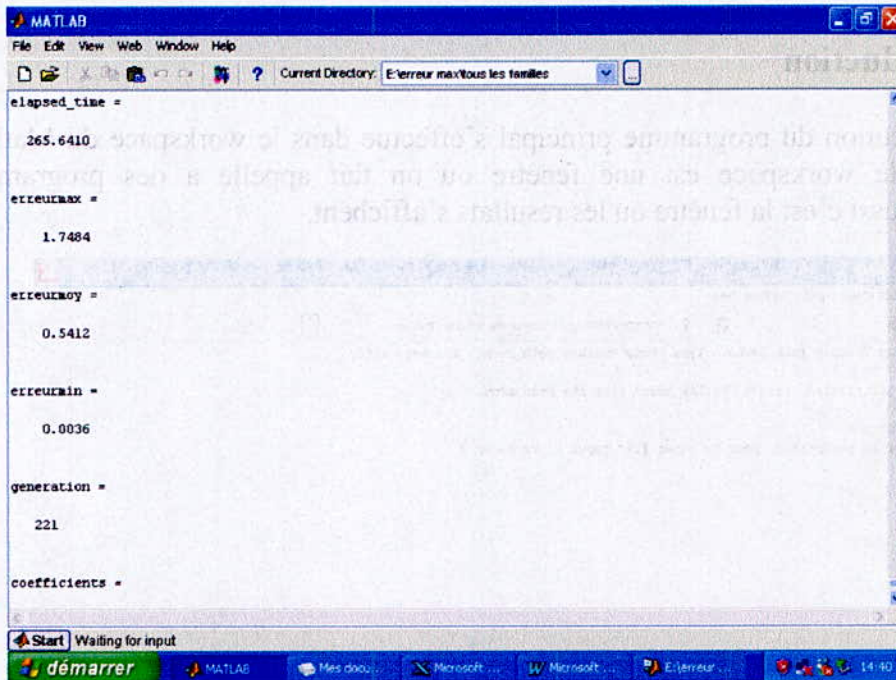


Figure IV.2

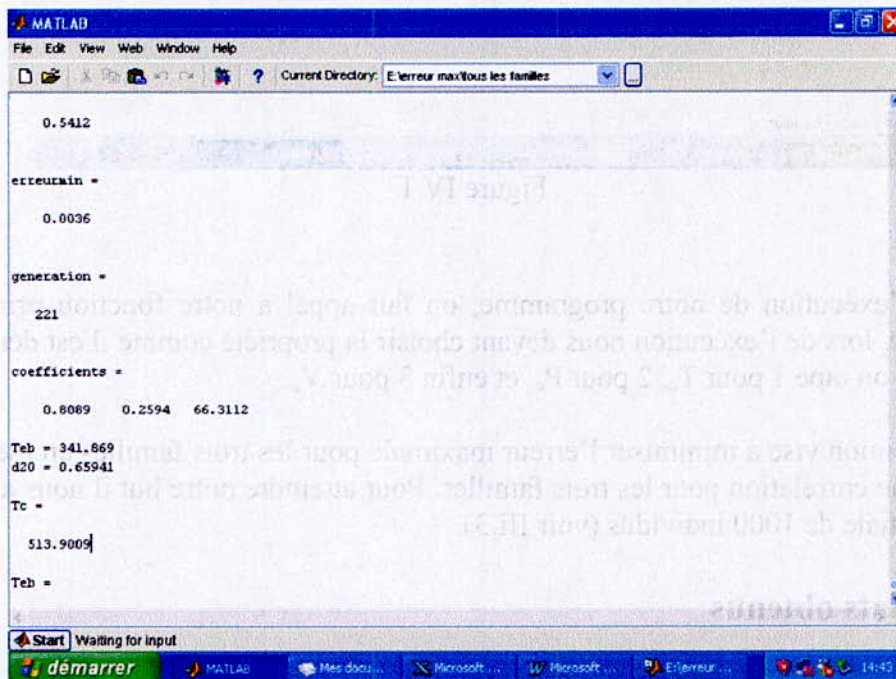


Figure IV.3

Le modèle obtenu par les AG est le suivant :

$$T_c = 0.8089 T_{eb} + 0.2594 10^3 d_{20} + 66.3112$$

L'application du programme proposé consiste à donner la température d'ébullition et la densité du corps choisi pour obtenir directement sa température critique. Comme le montre l'exemple de la Figure IV.3 où pour une entrée en température d'ébullition et une densité correspondante à la paraffine C<sub>6</sub> et pour lequel nous pouvons lire la valeur de sa température critique :  $T_c = 513.909 \text{ K}$ . L'erreur calculée pour une valeur expérimentale de 507.5 K est de 1.08 %.

Tous les résultats sont données en Annexe.II.1.

Nous avons représenté dans la Figure IV.4 les valeurs expérimentales et celle trouvées par le modèle proposé par les AG. Il est clair qu'il n'existe pas une grande différence entre ces deux valeurs. Nous avons aussi représenté l'erreur pour chaque valeur de  $T_c$  (Figure IV.5).

Figure IV.4 : Comparaison entre les valeurs calculées et expérimentales de la température critique

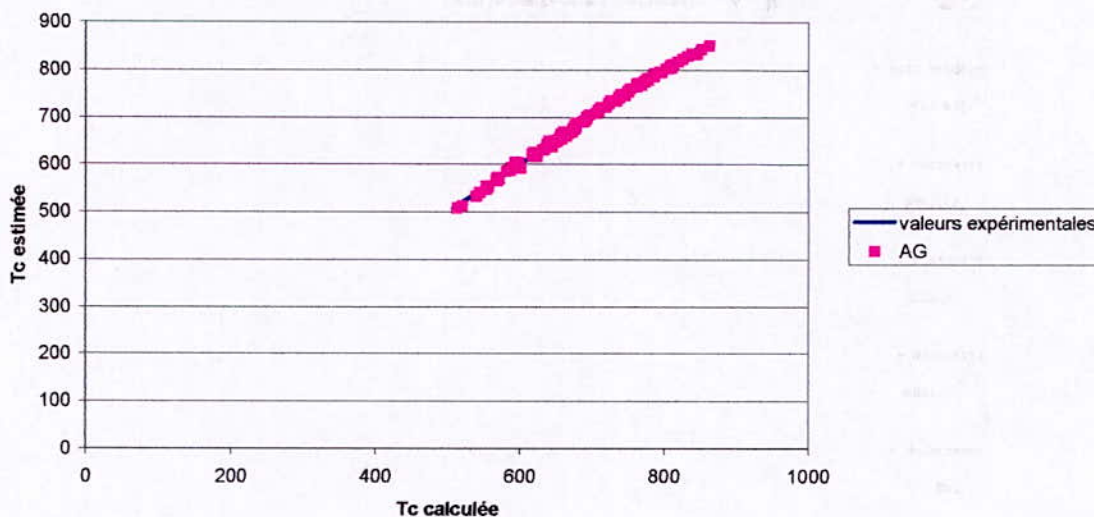
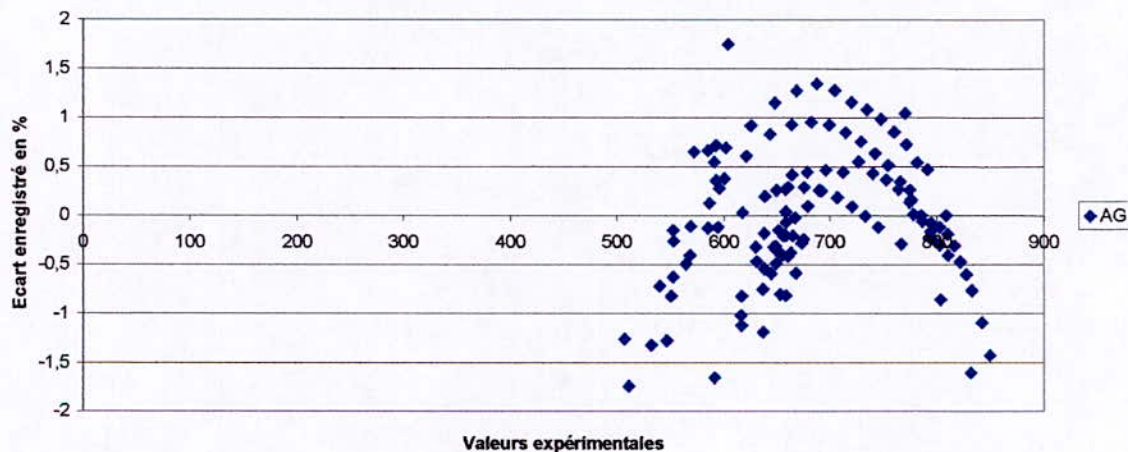


Figure IV.5 : Ecart enregistré pour la température critique



**Remarque :**

Il est possible d'améliorer les résultats obtenus en changeant seulement le modèle choisi dans le Chapitre III.1, et un outil plus performant.

**IV.2.2. Pression critique**

Après plusieurs essais nous avons obtenu une erreur maximale inférieure à 13.2% pour une population initiale de 1000 individus, le programme converge après un temps d'exécution ne dépassant les six minutes, pour un nombre de génération égale à 247 (Figure IV.6).

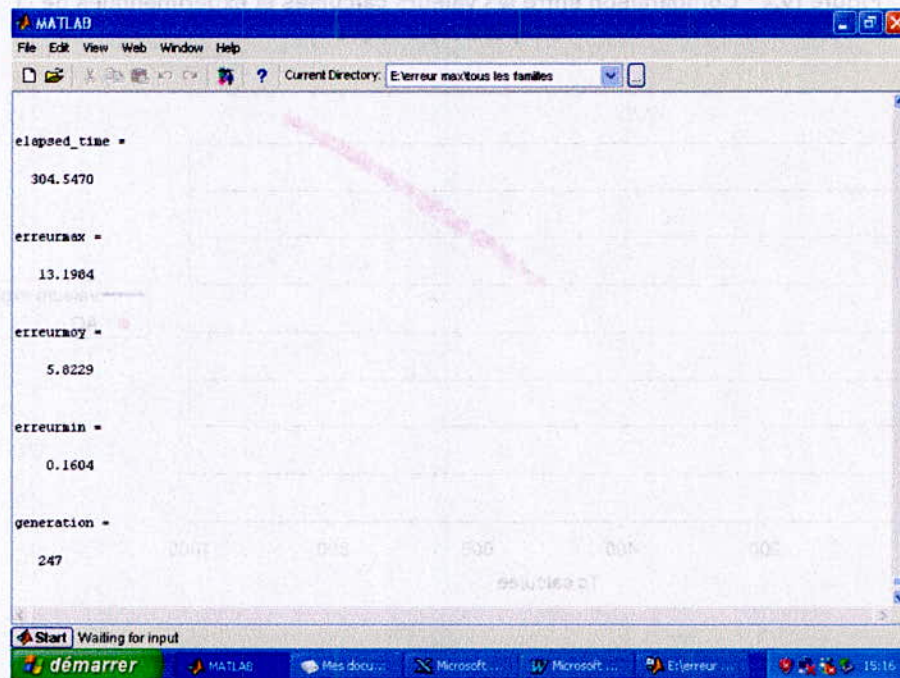


Figure IV.6



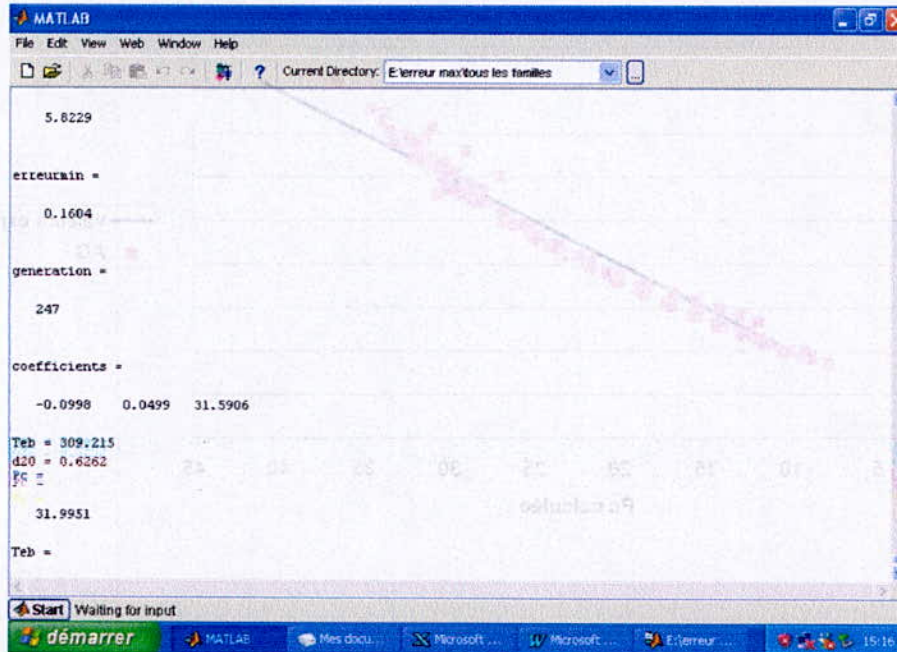


Figure IV.7

Les résultats obtenus sont comme suite :

L'erreur maximale = 13.1984 %  
 L'erreur moyenne = 5.8229 %  
 L'erreur minimale = 0.1604 %

Le modèle obtenu par les AG est le suivant :

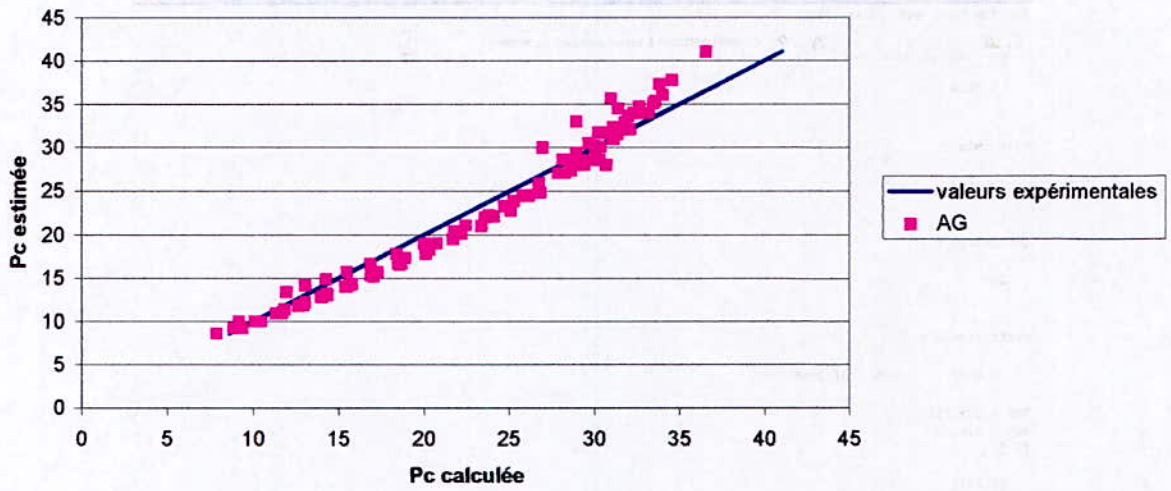
$$P_c = -0.0998 T_{eb} + 0.0499 \cdot 10^3 d_{20} + 31.5906$$

L'application du programme proposé consiste à donner la température d'ébullition et la densité du corps choisi pour obtenir directement sa pression critique. Comme le montre l'exemple de la Figure IV.7 où pour une entrée en température d'ébullition et une densité correspondante à la paraffine C<sub>5</sub> et pour lequel nous pouvons lire la valeur de sa pression critique : P<sub>c</sub> = 31.9951 bar . L'erreur calculée pour une valeur expérimentale de 33.65 bar est de 4.9179 % .

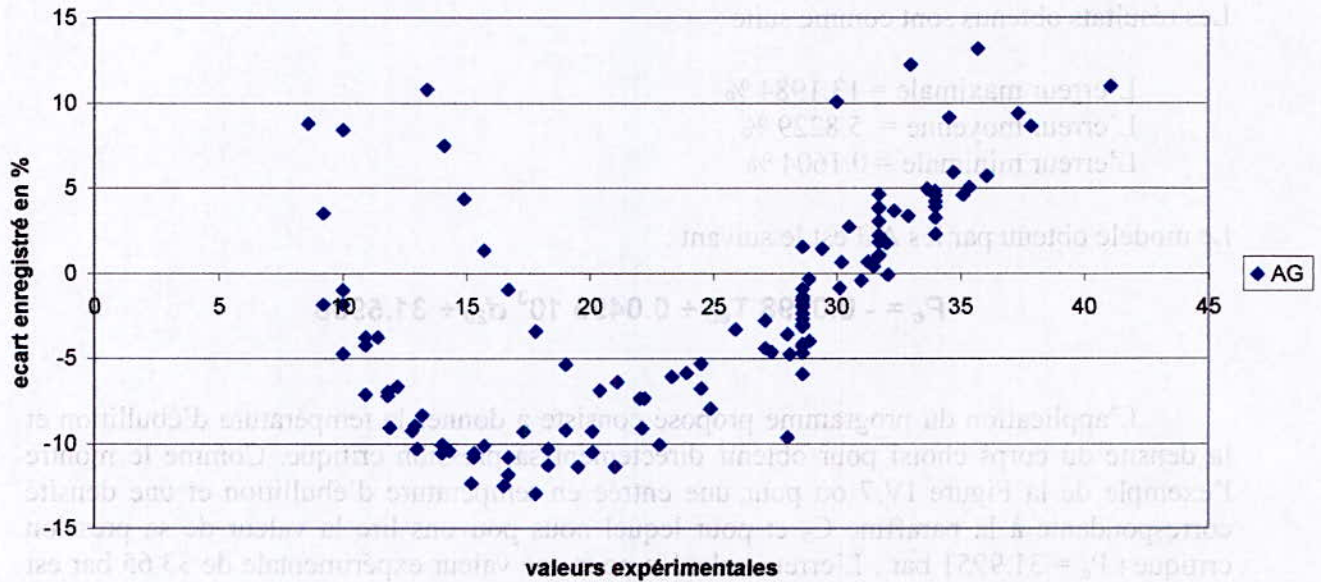
Tous les résultats sont données en Annexe.II.2.

Nous avons reporté les valeurs expérimentales et les valeurs trouvées par le modèle proposé dans un même graphe pour comparer les résultats obtenus avec les valeurs expérimentales (Figure IV.8), aussi nous avons représenté sur la Figure IV.9 l'erreur de chaque valeur de P<sub>c</sub>.

**Figure IV.8 : Comparaison des valeurs calculées et expérimentales de la pression critique**



**Figure IV.9 : Ecart enregistré pour les valeurs de la pression critique**



### IV.2.3. Volume critique

Pour l'exécution de cette propriété il faut bien sûr taper 3, le programme est lancé, et exécute les différents opérateurs présentés dans le chapitre III, jusqu'à atteindre l'erreur fixé au départ. En ce qui concerne l'erreur nous l'avons fixé suite à de nombreux essais à 9.8 %, ce qui veut dire que notre programme va s'arrêter quand il aura atteint une erreur maximale inférieure à 9.8 %. Après un temps moins de trois minutes le programme nous a donné ces résultats (Figure IV.10).

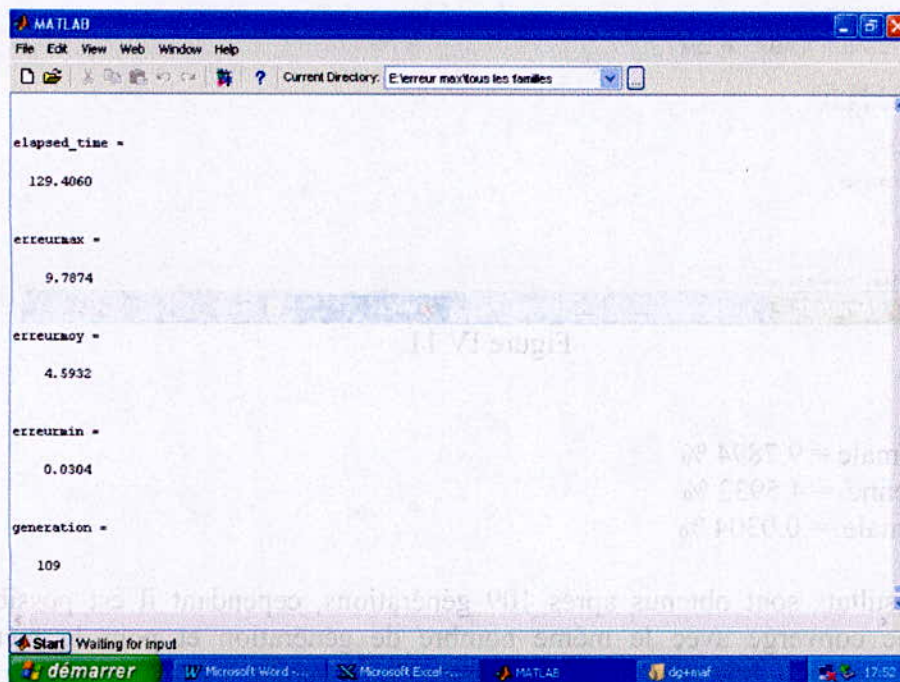


Figure IV.10



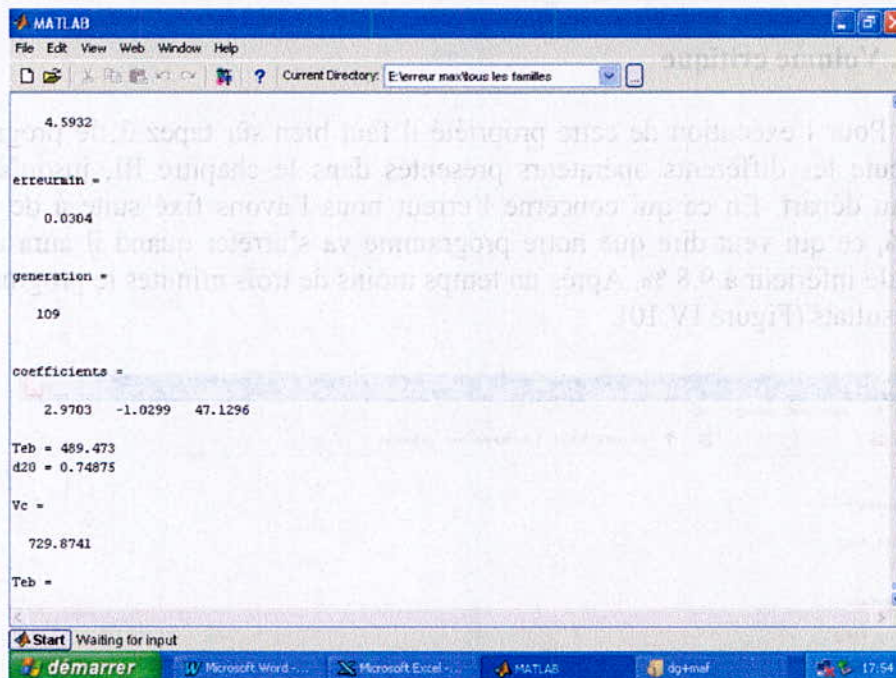


Figure IV.11

L'erreur maximale = 9.7894 %

L'erreur moyenne = 4.5932 %

L'erreur minimale = 0.0304 %

Ces résultats sont obtenus après 109 générations, cependant 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 des AG fonctionnent d'une façon aléatoire.

Le modèle obtenu est le suivant :

$$V_c = 2,9703 T_{eb} - 1,0299 \cdot 10^3 d_{20} + 47,1296$$

L'application du programme proposé consiste à donner la température d'ébullition et la densité du corps choisi pour obtenir directement son volume critique. Comme le montre l'exemple de la Figure IV.11 où pour une entrée en température d'ébullition et une densité correspondante à la paraffine C<sub>12</sub> et pour lequel nous pouvons lire la valeur de son volume critique :  $V_c = 729.8741 \text{ cm}^3 / \text{mol}$ . L'erreur calculée pour une valeur expérimentale de  $700 \text{ cm}^3 / \text{mol}$  est de 4.26 %.

Tous les résultats trouvés par cette corrélation proposée par les AG sont dans le tableau Annexe.II.3.

Nous avons tracé les valeurs expérimentales et les valeurs trouvées par le modèle présenté dans un même graphe afin de comparer nos résultats avec les valeurs expérimentales (Figure IV.12), la Figure IV.13 nous donne l'erreur pour chaque valeur expérimentale de  $V_c$ .

Figure IV.12 : Comparaison entre les valeurs calculés et expérimentales du volume critique

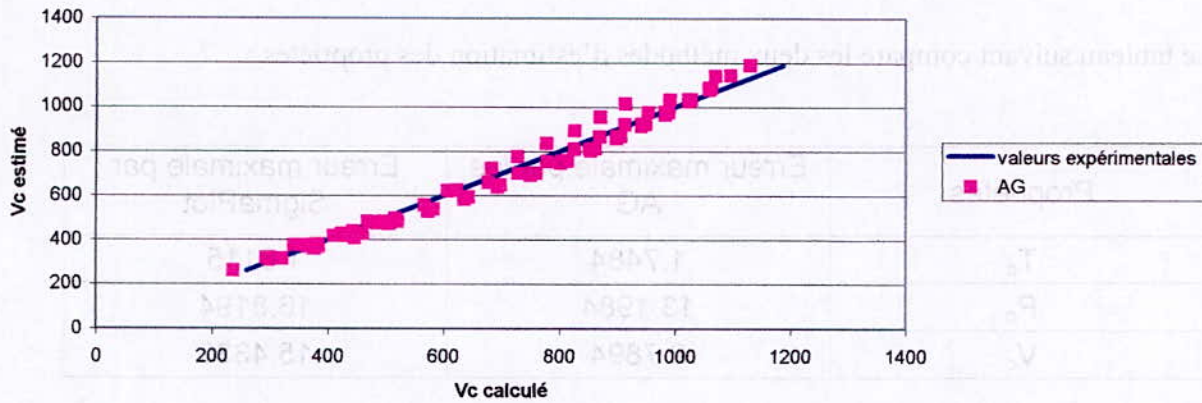
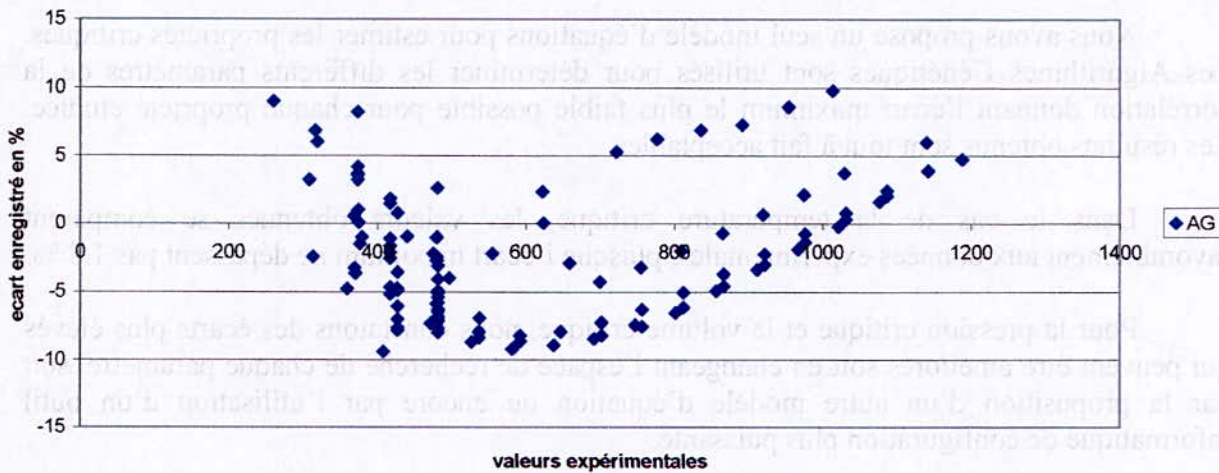


Figure IV.13 : Ecarts enregistrés pour le volume critique



### IV.3. Etude comparative avec le logiciel SigmaPlot

SigmaPlot est un logiciel qui donne les coefficients d'un modèle proposé par l'ingénieur, pour estimer et obtenir les caractéristiques désirés, nous avons voulu faire une étude comparative avec ce logiciel utilisé par les chercheurs au niveau du Laboratoire de Valorisation des Energies Fossiles, et ceci dans le but de valoriser le programme proposé.

Pour le modèle proposé, SigmaPlot donne pour chaque propriétés les corrélations suivantes :

$$T_c = 0,8143 T_{eb} + 0,2464 \cdot 10^3 d_{20} + 74,3572$$

$$P_c = -0,1077 T_{eb} + 0,0553 \cdot 10^3 d_{20} + 30,72$$

$$V_c = 2,7397 T_{eb} - 1,0461 \cdot 10^3 d_{20} + 147,8946$$

Le tableau suivant compare les deux méthodes d'estimation des propriétés :

Propriétés	Erreur maximale par les AG	Erreur maximale par SigmaPlot
$T_c$	1.7484	1.8115
$P_c$	13.1984	16,8194
$V_c$	9.7894	15,4333

Il ressort un bon accord entre les différentes valeurs. Toutefois les résultats obtenus par les AG sont meilleurs que ceux obtenus par le logiciel SigmaPlot qui nécessite à chaque fois une réinitialisation des valeurs des coefficients du modèle.

#### IV.4. Discussions

Nous avons proposé un seul modèle d'équations pour estimer les propriétés critiques. Les Algorithmes Génétiques sont utilisés pour déterminer les différents paramètres de la corrélation 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, les valeurs obtenues se comparent favorablement aux données expérimentales, puisque l'écart maximum ne dépassent pas 1.8 %.

Pour la pression critique et le volume critique, nous constatons des écarts plus élevés qui 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.

Il est important de faire remarquer que pour chaque propriété critique étudiée, une seule banque de données englobant les trois familles (paraffine, naphtène et aromatique) est utilisée pour chaque entrée ( $T_{eb}$ ,  $d_{20}$  et propriétés critiques).

Le but visé est l'application de la méthode à un large domaine de température et au calcul des propriétés des mélanges complexes d'hydrocarbures tels que les fractions pétrolières contenant quelques centaines de constituants appartenant à plusieurs familles chimiques et dont la composition est souvent déterminée d'une façon approximative.

# Conclusion Générale

La connaissance des propriétés physiques et thermodynamiques du pétrole et des coupes qu'en découlent, pour un ingénieur, est d'une importance capitale dans la mise au point de nouveaux procédés tels que le dimensionnement des équipements d'une raffinerie, la séparation et la récupération des produits...

Devant l'impossibilité d'effectuer certaines mesures expérimentales, différentes études ont été réalisées par détermination indirecte de ces propriétés.

Notre travail a consisté à développer un nouveau modèle de calcul des caractéristiques des hydrocarbures basé sur les Algorithmes Génétiques. Nous avons étudié les propriétés critiques dont la mesure expérimentale nécessite un appareillage lourd et faisant appel à un personnel compétent.

Nous proposons donc une méthode simple et rapide qui évite ces inconvénients et donne des résultats ne dépassant pas les 13.2 % dans le cas le plus défavorable.

Cependant, ils peuvent être améliorés de plusieurs manières :

- Changer l'espace de recherche
- Choisir un modèle d'équation adéquat à chaque propriété
- Coder les paramètres d'entrée en binaire comme par exemple pour une température critique de 513.9 K pour laquelle  $T_{eb} = 341.87$  K et  $d_{20} = 0.6594$  que nous pouvons coder comme suit :

$T_{eb}$	1	1	0	1	0	1	0	1	1	0	1	0	
$d_{20}$	1	1	0	0	1	1	1	0	0	0	0	1	0

On fera remarquer que la valeur de la densité a été multipliée par 10000 et celle de température d'ébullition par 10, parce que le codage en binaire ne prend en considération que la partie entière de la valeur.

Tous les opérateurs (sélection, croisement, mutation) vont intervenir pour trouver la meilleure combinaison possible entre les chromosomes de  $T_{eb}$  et  $d_{20}$  pour donner un vecteur qui correspond à la valeur de  $T_c$  codée :

$T_c$	1	0	1	0	0	0	0	0	0	1	0	1	0
-------	---	---	---	---	---	---	---	---	---	---	---	---	---

Cette méthode peut être intéressante parce qu'elle n'utilise pas de modèle mathématique.

Une des perspectives à explorer, qui à notre avis, peut aussi donner de bons résultats est le couplage des Réseaux de neurones aux AG.

Nous proposons ainsi pour les recherches à venir d'étudier d'autres propriétés et d'appliquer cette nouvelle technique de caractérisation aux fraction pétrolières.

## REFERENCES BIBLIOGRAPHIQUE

- [1] N.Skander « Contribution à la détermination des propriétés physiques des mélange d'hydrocarbures et des fractions pétroliers par des méthodes de contribution de groupes »  
Thèse de Magister, Département de Génie Chimique, Ecole Nationale Polytechnique (1997).
- [2] F.Souahi « contribution à l'étude physico-chimique des pétroles et des fractions pétrolière ;comparaison des méthodes expérimentales et empiriques aux méthodes numériques élaborées »,Thèse de Doctorat d'Etat, Département de Génie Chimique, Ecole Nationale Polytechnique(1999).
- [3] RNA Lamouni. PFE 2003  
Ecole Nationale Polytechnique
- [4] S.Gori « caractérisation des fractions pétrolières par différentes approches utilisant l'intelligence artificielle », projet de fin d'études. département de Génie Chimique. Ecole Nationale Polytechnique(2004).
- [5] RNA Hachmaoui .magistère 2004.  
Ecole Nationale Polytechnique
- [6] RNA Mezioud .magistère 2004.  
Ecole Nationale Polytechnique
- [7] [http://www.eudil.fr/~vmagnin/coursag/algo\\_evo.html](http://www.eudil.fr/~vmagnin/coursag/algo_evo.html)
- [8] <http://www.ai.univ-paris8.fr/~renaud/publications/hthese/node24>.
- [9] <http://www.toulouse.inra.fr/centre/esr/CV/bontemps/WP/AlgoGene.html>
- [10] <http://pro.chemist.online.fr/cours/ag1.htm>
- [11] <http://www.sylbarth.com/ag.php>
- [12] <http://www.ai.univ-paris8.fr/~renaud/publications/hthese/node23.html>
- [13] <http://www.guill.net/index.php>
- [14] <http://www.vieartificielle.com/article/index.php?id=44>
- [15] <http://www.renard.org/alife/french/gavintr.html>

REFERENCES BIBLIOGRAPHIQUES

[16] <http://www.ai.univ-paris8.fr/~renaud/publications/hthese/node22.html>

[17] **C.E Chitour**, « Raffinage, Tome 1 et Tome 2 »  
Edition OPU.

[18] **Abdelhalm Mesbah** « Apprendre et Maîtriser MATLAB »,  
Edition Spinger.

[19] **P. Wuithier**, Raffinage et Génie Chimique  
Edition technip (1972).



# Annexe I

Banque de données

## I.1. La banque de données utilisée pour la température critique

$T_{eb}$ (K)	$d_{20}$	$T_c$ (K)
341,869	0,6594	507,5
371,553	0,6837	540,3
398,83	0,7026	568,83
423,968	0,7177	594,7
447,305	0,7301	617,7
469,08	0,7402	639
489,473	0,7488	658,2
508,616	0,7561	675,8
526,727	0,7626	691,9
543,835	0,7684	706,8
560,014	0,7735	720,6
575,17	0,7780	733,3
589,4	0,7828	745,2
616,9	0,7843	766,5
664,3	0,7990	804
704,4	0,8066	832
322,37	0,7457	511,6
344,98	0,7497	532,79
354,429	0,7785	553,5
374,09	0,7694	572,19
391,951	0,8110	604,3
364,99	0,7514	553
372,65	0,7712	565
363,9	0,7448	553
364,7	0,7488	551
360,96	0,7545	547
376,59	0,7680	569,5
392,74	0,7809	585,8
397,48	0,7828	593
396,68	0,7660	586,6
404,97	0,7879	602,4
404,099	0,7763	596,2
402,89	0,7962	601
394,51	0,7847	593,2
392,51	0,7626	585,5
396,58	0,7760	591,6
429,865	0,7936	625,6
429,75	0,7846	621,2
454,09	0,7992	648,4
453,6	0,7912	643,8
476,8	0,8037	669
476	0,7965	664,1
497,8	0,8076	687,7
497	0,8010	682,6

516,7	0,8048	699,5
518	0,8110	704,7
536,9	0,8149	720,3
535,2	0,8081	714,9
552,526	0,8110	729,1
554,7	0,8160	734,6
570,77	0,8186	747,7
568,9	0,8135	742,2
584,3	0,8157	754,4
586,2	0,8210	759,9
600,8	0,8220	771,1
599	0,8178	765,6
613	0,8196	776
614,7	0,8240	781,5
627	0,8250	791,2
626	0,8213	785,7
639	0,8228	794,7
652	0,8281	808,5
383,762	0,8668	591,8
417,579	0,8801	630,33
409,32	0,8669	617,2
411,509	0,8610	616,23
412,27	0,8642	617,05
442,528	0,8758	649,17
425,561	0,8617	631,1
432,35	0,8615	638,32
437,893	0,8651	637,35
449,267	0,8944	664,47
438,33	0,8806	647,2
434,48	0,8645	637,2
435,166	0,8611	638,2
469,95	0,8875	675
442,269	0,8665	645
450,25	0,8573	651,9
456,42	0,8602	660,53
456,902	0,8620	659
456,573	0,8800	658,5
446,455	0,8621	649
454,252	0,8639	655,2
478,19	0,9052	689,7
471,15	0,8903	679
451,3	0,8766	653,5
448,2	0,8610	649
445,909	0,8532	650
461,35	0,8763	665,4
467,06	0,8921	673,6
456,73	0,8648	658,7
462,63	0,8745	667,2

458,08	0,8736	660,7
455,13	0,8609	656,4
456,53	0,8584	658,5
459,98	0,8772	663,4
463,16	0,8904	668
478	0,8591	678,8
499	0,8602	696,3
537,42	1,0059	770
520	0,8567	712,3
537	0,8567	726,9
556,37	0,9866	774,3
555	0,8558	740,4
567,44	0,9750	778,3
571	0,8554	752,8
581,04	0,9659	786,4
586	0,8553	764,2
594,18	0,9564	794,4
600,4	0,8551	774,8
607	0,9498	802,7
614	0,8550	784,7
619	0,9428	810
627	0,8549	793,9
630	0,9373	816,5
639	0,8548	802,5
640	0,9325	822
650	0,8547	810,5
650	0,9283	827,9
659	0,9245	832,8
676	0,9186	842,2
691	0,9135	850,1

### I.2. La banque de données utilisée pour la pression critique

$T_{eb}$ (K)	$d_{20}$	$P_c$ (bar)
309,215	0,6262	33,65
341,869	0,6594	30,12
371,553	0,6836	27,36
398,83	0,7025	24,87
423,968	0,7177	22,8
447,305	0,7301	21
469,08	0,7402	19,5
489,473	0,7487	17,8
508,616	0,7561	16,5
526,727	0,7625	15,2
543,835	0,7683	14,2
560,014	0,7735	13,2

575,17	0,778	12,2
589,4	0,7828	11,4
616,9	0,7843	10
344,98	0,7496	37,84
374,09	0,7693	34,71
364,99	0,7514	34
372,65	0,7712	34
363,9	0,7447	34
364,7	0,7487	34
360,96	0,7544	34
376,59	0,768	33,97
424,313	0,8361	35,7
392,74	0,7809	31,7
397,48	0,7828	31,7
396,68	0,7659	31,7
404,97	0,7878	31,7
404,099	0,7763	30,2
402,89	0,7962	31,7
394,51	0,7846	31,7
392,51	0,7625	31,7
396,58	0,7759	31,7
451,5	0,8502	33
429,865	0,7935	27,1
429,75	0,7846	27,1
475	0,8575	30
454,09	0,7992	24,5
453,6	0,7912	24,5
476,8	0,8037	22,1
476	0,7965	22,2
497,8	0,8076	20,1
497	0,801	20,1
516,7	0,8048	18,3
518	0,811	18,3
536,9	0,8149	16,7
535,2	0,8081	16,7
552,526	0,8109	15,3
554,7	0,816	15,3
570,77	0,8186	14
568,9	0,8135	14
584,3	0,8157	12,8
586,2	0,821	12,9
600,8	0,822	11,8
599	0,8178	11,8
613	0,8196	10,9
614,7	0,824	10,9
627	0,825	10
626	0,8213	10
639	0,8228	9,2

652	0,8281	8,6
383,762	0,8668	41,06
417,579	0,880	37,34
409,32	0,8669	36,06
411,509	0,861	35,11
412,27	0,8642	35,36
442,528	0,8758	32,32
425,561	0,8617	32,09
432,35	0,8615	32
437,893	0,8651	31,27
449,267	0,8944	34,54
438,33	0,8806	32,9
434,48	0,8645	31,5
435,166	0,8611	31
469,95	0,8875	29,4
442,269	0,8665	28
450,25	0,8573	28,1
456,42	0,8601	28,87
456,902	0,8619	28
456,573	0,8799	28,6
446,455	0,8620	28,9
454,252	0,8639	28,6
478,19	0,9052	28,6
471,15	0,8903	28,6
451,3	0,8766	28,6
448,2	0,861	28,6
445,909	0,8532	30,5
461,35	0,8763	28,6
467,06	0,8921	28,6
456,73	0,8648	28,6
462,63	0,8745	28,6
458,08	0,8736	28,6
455,13	0,8609	28,6
456,53	0,8584	28,6
459,98	0,8772	28,6
463,16	0,8904	28,6
478	0,8591	25,9
499	0,8602	23,3
537,42	1,0059	28,6
520	0,8567	21,1
537	0,8567	19
556,37	0,9866	23,9
555	0,8558	17,3
567,44	0,975	22
571	0,8554	15,7
581,04	0,9659	20,4
586	0,8553	14,3
594,18	0,9564	19

600,4	0,8551	13
607	0,9498	17,8
614	0,855	11,9
619	0,9428	16,7
627	0,8549	10,9
630	0,9373	15,7
639	0,8548	10
640	0,9325	14,9
650	0,8547	9,2
650	0,9283	14,1
659	0,9245	13,4

### I.3. La banque de données utilisée pour le volume critique

$T_{eb}$	$d_{20}$	$V_c$ (cm <sup>3</sup> /mol)
309,215	0,6262	313
341,869	0,6594	370
371,553	0,6837	426,4
398,83	0,7026	480
423,968	0,7177	555
447,305	0,7301	624
469,08	0,7402	659
489,473	0,7488	700
508,616	0,7561	755
526,727	0,7626	810
543,835	0,7684	865
560,014	0,7735	920
575,17	0,7780	975
589,4	0,7828	1030
616,9	0,7843	1140
322,37	0,7457	260
344,98	0,7497	319
354,429	0,7785	308
374,09	0,7694	368
391,951	0,8110	359
364,99	0,7514	373
372,65	0,7712	373
363,9	0,7448	373
364,7	0,7488	373
360,96	0,7545	373
376,59	0,7680	375
424,313	0,8361	408
392,74	0,7809	417
397,48	0,7828	417
396,68	0,7660	417
404,97	0,7879	417

404,099	0,7763	427
402,89	0,7962	417
394,51	0,7847	417
392,51	0,7626	417
396,58	0,7760	417
429,865	0,7936	472
429,75	0,7846	482
454,09	0,7992	527
453,6	0,7912	537
476,8	0,8037	582
476	0,7965	592
497,8	0,8076	637
497	0,8010	647
516,7	0,8048	702
518	0,8110	692
536,9	0,8149	747
535,2	0,8081	757
552,526	0,8110	812
554,7	0,8160	802
570,77	0,8186	857
568,9	0,8135	867
584,3	0,8157	922
586,2	0,8210	912
600,8	0,8220	967
599	0,8178	977
613	0,8196	1032
614,7	0,8240	1022
627	0,8250	1077
626	0,8213	1087
639	0,8228	1142
652	0,8281	1187
383,762	0,8668	316
417,579	0,8801	369
409,32	0,8669	374
411,509	0,8610	379
412,27	0,8642	376
442,528	0,8758	427
425,561	0,8617	427
432,35	0,8615	440
437,893	0,8651	427
449,267	0,8944	427
438,33	0,8806	427
434,48	0,8645	427
435,166	0,8611	427
442,269	0,8665	481
450,25	0,8573	482
456,42	0,8602	497
456,902	0,8620	482



456,573	0,8800	482
446,455	0,8621	475
454,252	0,8639	482
451,3	0,8766	480
448,2	0,8610	480
445,909	0,8532	480
461,35	0,8763	482
467,06	0,8921	482
456,73	0,8648	482
462,63	0,8745	482
458,08	0,8736	482
455,13	0,8609	482
456,53	0,8584	482
459,98	0,8772	482
463,16	0,8904	482
478	0,8591	537
499	0,8602	592
537,42	1,0059	622
528,45	1,0113	538
537	0,8567	702
556,37	0,9866	722
555	0,8558	757
567,44	0,9750	777
571	0,8554	812
581,04	0,9659	836
586	0,8553	867
594,18	0,9564	892
600,4	0,8551	922
607	0,9498	954
614	0,8550	977
619	0,9428	1014
627	0,8549	1032
639	0,8548	1087
650	0,8547	1142

# Annexe II

Résultats trouvés

## II.1. Résultats trouvés pour la température critique

$T_{eb}$ (K)	$d_{20}$	$T_c$ (K)	l'erreur (%)
341,869	0,6594	513,9	-1,2611
371,553	0,6837	544,2096	-0,7236
398,83	0,7026	571,1689	-0,4112
423,968	0,7177	595,4355	-0,1237
447,305	0,7301	617,5293	0,0276
469,08	0,7402	637,7579	0,1944
489,473	0,7488	656,4717	0,2626
508,616	0,7561	673,863	0,2866
526,727	0,7626	690,1861	0,2477
543,835	0,7684	705,5371	0,1787
560,014	0,7735	719,9602	0,0888
575,17	0,7780	733,3794	-0,0108
589,4	0,7828	746,1352	-0,1255
616,9	0,7843	768,769	-0,2960
664,3	0,7990	810,9241	-0,8612
704,4	0,8066	845,3324	-1,6025
322,37	0,7457	520,5212	-1,7438
344,98	0,7497	539,8325	-1,3218
354,429	0,7785	554,9595	-0,2637
374,09	0,7694	568,482	0,6480
391,951	0,8110	593,7338	1,7485
364,99	0,7514	556,4726	-0,6279
372,65	0,7712	567,8048	-0,4964
363,9	0,7448	553,8607	-0,1556
364,7	0,7488	555,5506	-0,8259
360,96	0,7545	553,9987	-1,2795
376,59	0,7680	570,1541	-0,1148
392,74	0,7809	586,5692	-0,1313
397,48	0,7828	590,8989	0,3543
396,68	0,7660	585,8835	0,1222
404,97	0,7879	598,2649	0,6864
404,099	0,7763	594,5669	0,2739
402,89	0,7962	598,7536	0,3738
394,51	0,7847	588,9737	0,7125
392,51	0,7626	581,6258	0,6617
396,58	0,7760	588,3888	0,5428
429,865	0,7936	619,8785	0,9146
429,75	0,7846	617,4612	0,6019
454,09	0,7992	640,9371	1,1510
453,6	0,7912	638,4655	0,8286
476,8	0,8037	660,4745	1,2744
476	0,7965	657,9597	0,9246
497,8	0,8076	678,4731	1,3417
497	0,8010	676,1139	0,9502

516,7	0,8048	693,035	0,9242
518	0,8110	695,6948	1,2779
536,9	0,8149	711,9947	1,1530
535,2	0,8081	708,8556	0,8455
552,526	0,8110	723,6151	0,7523
554,7	0,8160	726,6784	1,0784
570,77	0,8186	740,3519	0,9828
568,9	0,8135	737,5163	0,6311
584,3	0,8157	750,5441	0,5111
586,2	0,8210	753,4558	0,8480
600,8	0,8220	765,5251	0,7230
599	0,8178	762,9796	0,3423
613	0,8196	774,7711	0,1584
614,7	0,8240	777,2876	0,5390
627	0,8250	787,4965	0,4681
626	0,8213	785,7278	-0,0035
639	0,8228	796,6326	-0,2432
652	0,8281	808,5231	-0,0029
383,762	0,8668	601,5946	-1,6550
417,579	0,8801	632,3888	-0,3266
409,32	0,8669	622,2892	-0,8246
411,509	0,8610	622,5242	-1,0214
412,27	0,8642	623,9699	-1,1214
442,528	0,8758	651,4546	-0,3519
425,561	0,8617	634,0725	-0,4710
432,35	0,8615	639,5122	-0,1868
437,893	0,8651	644,9298	-1,1893
449,267	0,8944	661,7306	0,4123
438,33	0,8806	649,304	-0,3251
434,48	0,8645	642,0134	-0,7554
435,166	0,8611	641,6863	-0,5463
469,95	0,8875	676,6713	-0,2476
442,269	0,8665	648,8327	-0,5942
450,25	0,8573	652,902	-0,1537
456,42	0,8602	658,6374	0,2865
456,902	0,8620	659,4917	-0,0746
456,573	0,8800	663,8947	-0,8192
446,455	0,8621	651,0696	-0,3189
454,252	0,8639	657,8617	-0,4062
478,19	0,9052	687,928	0,2569
471,15	0,8903	678,3683	0,0930
451,3	0,8766	658,7578	-0,8046
448,2	0,8610	652,2036	-0,4936
445,909	0,8532	648,3297	0,2570
461,35	0,8763	666,8094	-0,2118
467,06	0,8921	675,5268	-0,2860
456,73	0,8648	660,0892	-0,2109
462,63	0,8745	667,3779	-0,0267

458,08	0,8736	663,464	-0,4183
455,13	0,8609	657,7833	-0,2107
456,53	0,8584	658,2673	0,0353
459,98	0,8772	665,9347	-0,3821
463,16	0,8904	671,9311	-0,5885
478	0,8591	675,8159	0,4396
499	0,8602	693,0882	0,4613
537,42	1,0059	761,9607	1,0441
520	0,8567	709,1672	0,4398
537	0,8567	722,9185	0,5477
556,37	0,9866	772,2829	0,2605
555	0,8558	737,2452	0,4261
567,44	0,9750	778,2284	0,0092
571	0,8554	750,0839	0,3608
581,04	0,9659	786,8689	-0,0596
586	0,8553	762,1914	0,2628
594,18	0,9564	795,0336	-0,0798
600,4	0,8551	773,7877	0,1307
607	0,9498	803,6916	-0,1235
614	0,8550	784,7628	-0,0080
619	0,9428	811,5826	-0,1954
627	0,8549	795,2526	-0,1704
630	0,9373	819,0538	-0,3128
639	0,8548	804,9334	-0,3032
640	0,9325	825,8977	-0,4742
650	0,8547	813,8054	-0,4078
650	0,9283	832,8972	-0,6036
659	0,9245	839,1916	-0,7675
676	0,9186	851,4124	-1,0939
691	0,9135	862,223	-1,4261

## II.2. Résultats trouvés pour la pression critique

$T_{eb}$ (K)	$d_{20}$	$P_c$ (bar)	l'erreur (%)
309,215	0,6262	31,9783	4,9678
341,869	0,6594	30,3766	-0,8520
371,553	0,6836	28,6257	-4,6262
398,83	0,7025	26,8451	-7,9417
423,968	0,7177	25,0928	-10,0562
447,305	0,7301	23,3825	-11,3455
469,08	0,7402	21,7123	-11,3456
489,473	0,7487	20,1038	-12,9428
508,616	0,7561	18,5601	-12,4855
526,727	0,7625	17,0744	-12,3322
543,835	0,7683	15,6580	-10,2678
560,014	0,7735	14,3003	-8,3360

575,17	0,778	13,0108	-6,6462
589,4	0,7828	11,8302	-3,7737
616,9	0,7843	9,1605	8,3945
344,98	0,7496	34,5706	8,6400
374,09	0,7693	32,6469	5,9436
364,99	0,7514	32,6609	3,9384
372,65	0,7712	32,8845	3,2809
363,9	0,7447	32,4369	4,5973
364,7	0,7487	32,5576	4,2422
360,96	0,7544	33,2143	2,3107
376,59	0,768	32,3301	4,8274
424,313	0,8361	30,9655	13,1984
392,74	0,7809	31,3630	1,0629
397,48	0,7828	30,9853	2,2545
396,68	0,7659	30,2248	4,6535
404,97	0,7878	30,4893	3,8192
404,099	0,7763	30,0003	0,6610
402,89	0,7962	31,1145	1,8468
394,51	0,7846	31,3735	1,0299
392,51	0,7625	30,4708	3,8775
396,58	0,7759	30,7323	3,0526
451,5	0,8502	28,9558	12,2549
429,865	0,7935	28,2887	-4,3864
429,75	0,7846	27,8530	-2,7789
475	0,8575	26,9748	10,0838
454,09	0,7992	26,1524	-6,7449
453,6	0,7912	25,8022	-5,3151
476,8	0,8037	24,1105	-9,0977
476	0,7965	23,8311	-7,3475
497,8	0,8076	22,2094	-10,4945
497	0,801	21,9599	-9,2532
516,7	0,8048	20,1834	-10,2921
518	0,811	20,3631	-11,2738
536,9	0,8149	18,6714	-11,8053
535,2	0,8081	18,5018	-10,7894
552,526	0,8109	16,9159	-10,5615
554,7	0,816	16,9499	-10,7839
570,77	0,8186	15,4758	-10,5421
568,9	0,8135	15,4080	-10,0574
584,3	0,8157	13,9808	-9,2257
586,2	0,821	14,0557	-8,9592
600,8	0,822	12,6485	-7,1912
599	0,8178	12,6186	-6,9375
613	0,8196	11,3112	-3,7728
614,7	0,824	11,3611	-4,2306
627	0,825	10,1835	-1,8350
626	0,8213	10,0986	-0,9867
639	0,8228	8,8761	3,5204

652	0,8281	7,8431	8,8001
383,762	0,8668	36,5464	10,9925
417,579	0,8801	33,8332	9,3915
409,32	0,8669	33,9997	5,7133
411,509	0,861	33,4859	4,6257
412,27	0,8642	33,5696	5,0633
442,528	0,8758	31,1287	3,6859
425,561	0,8617	32,1184	-0,0886
432,35	0,8615	31,4309	1,7784
437,893	0,8651	31,0573	0,6800
449,267	0,8944	31,3843	9,1363
438,33	0,8806	31,7872	3,3824
434,48	0,8645	31,368	0,4189
435,166	0,8611	31,1299	-0,4191
469,95	0,8875	28,9758	1,4427
442,269	0,8665	30,6905	-9,6089
450,25	0,8573	29,4349	-4,7506
456,42	0,8601	28,9623	-0,3199
456,902	0,8619	29,0035	-3,5842
456,573	0,8799	29,9346	-4,6665
446,455	0,8620	30,0516	-3,9851
454,252	0,8639	29,3668	-2,6813
478,19	0,9052	29,0367	-1,5270
471,15	0,8903	28,9958	-1,3839
451,3	0,8766	30,2932	-5,9203
448,2	0,861	29,8241	-4,2802
445,909	0,8532	29,6640	2,7408
461,35	0,8763	29,2752	-2,3610
467,06	0,8921	29,4938	-3,1252
456,73	0,8648	29,1624	-1,9667
462,63	0,8745	29,0576	-1,6003
458,08	0,8736	29,4668	-3,0310
455,13	0,8609	29,1275	-1,8445
456,53	0,8584	28,8630	-0,9198
459,98	0,8772	29,4568	-2,9961
463,16	0,8904	29,7981	-4,1895
478	0,8591	26,7552	-3,3023
499	0,8602	24,7143	-6,0703
537,42	1,0059	28,1504	1,5717
520	0,8567	22,4439	-6,3693
537	0,8567	20,7473	-9,1965
556,37	0,9866	25,2962	-5,8419
555	0,8558	18,9060	-9,2834
567,44	0,975	23,6125	-7,3299
571	0,8554	17,2892	-10,1227
581,04	0,9659	21,8012	-6,8687
586	0,8553	15,7872	-10,4005
594,18	0,9564	20,0157	-5,3463

600,4	0,8551	14,3401	-10,3090
607	0,9498	18,4070	-3,4102
614	0,855	12,9779	-9,0580
619	0,9428	16,8601	-0,9588
627	0,8549	11,6755	-7,1148
630	0,9373	15,4878	1,3511
639	0,8548	10,4729	-4,7292
640	0,9325	14,2503	4,3601
650	0,8547	9,3701	-1,8492
650	0,9283	13,0427	7,4981
659	0,9245	11,9549	10,7840

### II.3. Résultats trouvés pour le volume critique

Teb (K)	d <sub>20</sub>	Vc (cm <sup>3</sup> /mol)	l'erreur (%)
309,215	0,6262	320,6675	-2,4497
341,869	0,6594	383,4567	-3,6370
371,553	0,6836	446,6211	-4,7423
398,83	0,7025	508,2078	-5,8766
423,968	0,7177	567,2619	-2,2094
447,305	0,7301	623,8091	0,0306
469,08	0,7402	678,1059	-2,8992
489,473	0,7487	729,8736	-4,2677
508,616	0,7561	779,1643	-3,2006
526,727	0,7625	826,3166	-2,0144
543,835	0,7683	871,1281	-0,7085
560,014	0,7735	913,8806	0,6651
575,17	0,778	954,2949	2,1236
589,4	0,7828	991,6187	3,7263
616,9	0,7843	1071,757	5,9862
322,37	0,7457	236,6276	8,9894
344,98	0,7496	299,7283	6,0413
354,429	0,7785	298,082	3,2201
374,09	0,7693	365,9356	0,5610
391,951	0,811	376,0928	-4,7612
364,99	0,7514	357,3616	4,1926
372,65	0,7712	359,7221	3,5598
363,9	0,7447	360,9934	3,2189
364,7	0,7487	359,2295	3,6918
360,96	0,7544	342,2707	8,2384
376,59	0,768	374,7517	0,0662
424,313	0,8361	446,3671	-9,4037
392,74	0,7809	409,4157	1,8188
397,48	0,7828	421,5278	-1,0858
396,68	0,7659	436,4951	-4,6751
404,97	0,7878	438,5847	-5,1762



404,099	0,7763	447,8826	-4,8905
402,89	0,7962	423,7862	-1,6274
394,51	0,7846	410,811	1,4842
392,51	0,7625	427,6209	-2,5470
396,58	0,7759	425,93	-2,1415
429,865	0,7935	506,6702	-7,3454
429,75	0,7846	515,5565	-6,9619
454,09	0,7992	572,817	-8,6939
453,6	0,7912	579,6008	-7,9331
476,8	0,8037	635,638	-9,2162
476	0,7965	640,6771	-8,2225
497,8	0,8076	693,9977	-8,9478
497	0,801	698,4188	-7,9473
516,7	0,8048	753,0201	-7,2678
518	0,811	750,4961	-8,4532
536,9	0,8149	802,6182	-7,4455
535,2	0,8081	804,572	-6,2843
552,526	0,8109	853,0796	-5,0591
554,7	0,816	854,3566	-6,5283
570,77	0,8186	899,4116	-4,9488
568,9	0,8135	899,1096	-3,7035
584,3	0,8157	942,5865	-2,2328
586,2	0,821	942,7716	-3,3741
600,8	0,822	985,108	-1,8726
599	0,8178	984,0871	-0,7254
613	0,8196	1023,817	0,7929
614,7	0,824	1024,335	-0,2285
627	0,825	1059,84	1,5933
626	0,8213	1060,681	2,4213
639	0,8228	1097,75	3,8748
652	0,8281	1130,905	4,7258
383,762	0,8668	294,2594	6,8800
417,579	0,8801	381,0495	-3,2655
409,32	0,8669	370,0919	1,0449
411,509	0,861	382,6909	-0,9738
412,27	0,8642	381,6556	-1,5041
442,528	0,8758	459,5841	-7,6309
425,561	0,8617	423,7086	0,7708
432,35	0,8615	444,08	-0,9273
437,893	0,8651	456,8367	-6,9875
449,267	0,8944	460,4448	-7,8325
438,33	0,8806	442,1713	-3,5530
434,48	0,8645	447,317	-4,7581
435,166	0,8611	452,8563	-6,0553
442,269	0,8665	468,3929	2,6210
450,25	0,8573	501,5739	-4,0610
456,42	0,8601	516,9448	-4,0130
456,902	0,8619	516,533	-7,1645

456,573	0,8799	497,0176	-3,1157
446,455	0,8620	485,389	-2,1872
454,252	0,8639	506,6225	-5,1084
451,3	0,8766	484,8157	-1,0033
448,2	0,861	491,6742	-2,4321
445,909	0,8532	492,8921	-2,6859
461,35	0,8763	514,9761	-6,8415
467,06	0,8921	515,6641	-6,9843
456,73	0,8648	513,0972	-6,4517
462,63	0,8745	520,6319	-8,0149
458,08	0,8736	508,044	-5,4033
455,13	0,8609	512,3613	-6,2990
456,53	0,8584	519,0945	-7,6960
459,98	0,8772	509,9799	-5,8050
463,16	0,8904	505,8308	-4,9441
478	0,8591	582,1459	-8,4071
499	0,8602	643,3893	-8,6806
537,42	1,0059	607,4518	2,3389
528,45	1,0113	575,2468	-6,9232
537	0,8567	759,8654	-8,2429
556,37	0,9866	683,6161	5,3163
555	0,8558	814,2577	-7,5638
567,44	0,975	728,4441	6,2491
571	0,8554	862,1944	-6,1816
581,04	0,9659	778,2123	6,9124
586	0,8553	906,8519	-4,5965
594,18	0,9564	827,0261	7,2841
600,4	0,8551	949,8302	-3,0185
607	0,9498	871,9027	8,6056
614	0,855	990,3293	-1,3643
619	0,9428	914,7556	9,7874
627	0,8549	1029,046	0,2862
639	0,8548	1064,793	2,0430
650	0,8547	1097,569	3,8906