

وزارة الجامعات والبحث العلمي  
Ministère aux Universités et de la Recherche Scientifique

ECOLE NATIONALE POLYTECHNIQUE

المدرسة الوطنية المتعددة التقنيات  
المكتبة — BIBLIOTHEQUE  
Ecole Nationale Polytechnique

DEPARTEMENT G. INDUSTRIEL

# PROJET DE FIN D'ETUDES

## SUJET

SUR LE PROBLEME DU  
VOYAGEUR DE COMMERCE

Proposé par :

Etudié par :

Dirigé par :

M<sup>r</sup> Ouabdesselam

A. Skoudarli

M<sup>r</sup> Ouabdesselam

A. Talah

M<sup>r</sup> Sari

PROMOTION

Juillet 92

المدرسة الوطنية المتعددة التقنيات  
المكتبة — BIBLIOTHEQUE  
Ecole Nationale Polytechnique

*"Les mots et les formules sont de grands générateurs d'opinions et de croyances.  
Puissances redoutables, ils ont fait périr plus d'hommes que les canons."*

Gustave LEBON

DEDICACES

المدرسة الوطنية المتعددة التقنيات  
BIBLIOTHEQUE — المكتبة  
Ecole Nationale Polytechnique

A ma Mère,  
qui a tant souffert.

A Fadhila,  
qui a su me supporter.

A Kenza,  
sur laquelle je ne dirai rien.

A.Talah

A mes Parents,  
A mes frères,  
A mes soeurs,  
A tous mes amis.

A.Skoudarli

REMERCIEMENTS

المدسة الوطنية المتعددة التقنيات  
المكتبة — BIBLIOTHEQUE  
Ecole Nationale Polytechnique

Nous adressons nos remerciements les plus profonds à notre inconditionnel promoteur Mr OUABDESSELAM qui, malgré son âge, nous a donné le meilleur de lui-même. Et nous a honoré en acceptant de nous diriger dans notre travail. Qu'il trouve ici le respect des élèves envers leur maître.

Il serait impardonnable d'oublier notre infatigable co-promoteur Mr SARI, qui n'a pas ménagé de ses efforts pour nous donner des conseils qui permettent à des fourmis placées sur l'épaule d'un géant de voir plus loin que lui. Nous le remercions pour sa disponibilité et sa compréhension inlassables.

Nos remerciements s'adressent aussi à l'ensemble des enseignants qui nous ont patiemment formés durant toutes ces longues années, à ceux d'entre eux qui ont, aujourd'hui, l'ultime occasion de juger notre modeste travail, et enfin à ces camarades de promotion dont le contact a été agréable et très enrichissant.

Nous ne saurions terminer sans adresser nos vifs remerciements à Melle ABOUN pour la sympathique atmosphère du département Génie Industriel dans lequel nous avons vécu.

Nous remercions, enfin, tous ceux et celles qui ont, de loin ou de près, contribué à la réalisation de ce projet. Qu'ils trouvent ici notre profond respect.

SOMMAIRE

Pages

INTRODUCTION	1
I/PRESENTATION DU PROBLEME A ETUDIER.	3
I.1.Introduction	3
I.2.Apparition du problème du voyageur du commerce(TSP)	3
I.3.Enoncé du problème	4
I.4.Définition de l'optimisation combinatoire	4
I.5.Classification du problème	5
5.1.Théorie de la complexité	
5.2.Algorithmes et complexité	
1°/Algorithmes	
2°/Evaluation des algorithmes	
3°/Classification des problèmes	
II/APERCU SUR LES TRAVAUX DE RECHERCHE.	8
II.1.Les différentes formulations	8
II.2.Approches de résolution	12
2.1.Approches exactes	
2.2.Approches heuristiques	
1°/Procédures de construction de tournées	
2°/Procédures d'affinage	
3°/Procédures mixtes	
2.3.Autres méthodes	
1°/Méthodes de recherche avec tabous	
2°/Méthodes de "recuit simulé"	
II.3.Rappels sur la théorie des graphes	17
II.4.Conclusion	21
III/APPROCHE STATISTIQUE DU PROBLEME DU VOYAGEUR DE COMMERCE.	23
III.1.Introduction	23

III.2. Historique	23
III.3. Estimation de la solution du TSP par intervalle de confiance en utilisant la loi de Weibull	25
3.1. Enoncé du problème	
3.2. Définition de la loi de Weibull	
3.3. Procédure d'estimation	
III.4. Détermination de la moyenne et de la variance de la population	35
IV/DESCRIPTION ET MISE EN OEUVRE DE LA PROCEDURE	45
IV.1. Procédure générale	45
IV.2. Choix et description de l'heuristique	48
IV.3. Choix de l'échantillon	51
IV.4. Implémentation	51
IV.5. Résultats et interprétations	53
CONCLUSION ET SUGGESTIONS	65
BIBLIOGRAPHIE	68
ANNEXES	72

PREAMBULE

Notre étude porte sur "le problème du voyageur de commerce".

Dans le cadre d'une approche statistique de la distribution des différents circuits possibles, nous avons été amenés à retrouver des résultats de certaines recherches de Mr OUABDESSELAM. Dans le même esprit, nous avons construit un intervalle de confiance dans lequel peut se trouver, avec un certain risque, la solution du problème [6, 7, 8].

Par ailleurs, nous avons repris certains travaux de Mr SARI et nous avons pu ainsi comparer les résultats obtenus par ces différentes méthodes.

## INTRODUCTION

La classe des problèmes qui définit le champ de l'optimisation combinatoire semble triviale à première vue. Un coût (ou un poids) ayant été affecté à chaque élément d'un ensemble fini, il s'agit de choisir un "meilleur" élément de coût minimum (ou de poids maximum).

L'importance pratique de ce type de question a nécessité la création d'outils mathématiques et algorithmiques, au cours d'une période récente, comme un corpus de connaissances ayant à la fois une grande richesse théorique et une forte application pratique. Ils se sont développés dans le cadre d'un rapport dialectique, "théorie-pratique" intense.

Le mot combinatoire évoque, outre le caractère fini de l'ensemble objet de l'étude, le cardinal très grand de cet ensemble ; ou plutôt l'existence d'une structure qui, à partir d'un nombre limité d'éléments, permet d'engendrer une quantité astronomique de situations à comparer. Ainsi, pour le très célèbre et classique problème du voyageur de commerce, (TSP en anglais) le nombre de tournées (circuits Hamiltoniens dans sa définition en théorie des graphes) à comparer, quand il y a plus de 100 villes, dépasse le nombre estimé d'atomes dans l'univers. Ce nombre est égal à  $n!$  lorsqu'on est en présence d'un graphe à  $(n+1)$  sommets. A titre indicatif,  $10! = 3628800$ , et pour une augmentation d'une unité, seulement, dans le nombre de sommets, le nombre de circuits à comparer devient  $11! = 39916800$ .

Il est alors clair que la procédure triviale de recherche d'une solution par énumération de l'ensemble des cas, si elle est

théoriquement possible, conduit à des procédures d'une efficacité désastreuse, ce qui est évidemment inacceptable, compte tenu de l'origine concrète des problèmes étudiés et de l'objectif opérationnel des méthodes élaborées.

Pour cette raison, la plupart des chercheurs dans ce domaine se sont orientés, par la force des choses, vers des méthodes dites heuristiques ou approximatives.

Une fois ces méthodes mises en oeuvre, un autre problème, non moins important, apparaît. Il s'agit de l'analyse de ces heuristiques et des solutions qu'elles donnent. Plusieurs procédés sont alors proposés [annexe 3]. Parmi ceux-ci, l'analyse empirique des heuristiques, qui renferme, entre autres, l'estimation statistique de la solution optimale du TSP.

Notre travail rentre dans le cadre de cette dernière classe. il est structuré de la manière suivante:

- Introduction.
- Présentation du problème à étudier.
- Aperçu sur les travaux de recherche.
- Approche statistique du problème du voyageur de commerce.
- Mise en oeuvre des procédures.
- Conclusion.

I/PRESENTATION DU PROBLEME A ETUDIER :I.1. INTRODUCTION:

Avant d'aborder la complexité algorithmique du problème du voyageur du commerce, ces différentes modélisations et méthodes de résolution, nous allons donner un aperçu historique de ce problème.

I.2. APPARITION DU PROBLEME DU VOYAGEUR DE COMMERCE (P.V.C ou T.S.P):

La première utilisation du terme TSP (Traveling Salesman Problem) dans les cercles mathématiques pourrait être en 1931-32.

L'histoire semble ne pas avoir retenu le nom de la personne qui a, la première, introduit l'appellation de TSP dans les cercles mathématiques, mais il n'y a pas de doute que Merrill FLOOD ait le plus important rôle dans l'introduction du TSP au sein de cette communauté et celle de la recherche opérationnelle. John Williams incita Flood en 1948 à populariser le TSP chez RAND & CO.

Une autre raison de la popularité du TSP était ses intimes connexions avec les éminents sujets dans les problèmes combinatoires tels que les problèmes d'affectation, et les problèmes de transport en général.

Etudié ensuite par Dantzig, Fulkerson et Johnson, cité dans [12] en 1954, le TSP n'a cessé, jusqu'à nos jours, d'être un sujet de réflexion de plusieurs chercheurs.

réflexion de plusieurs chercheurs.

### I.3. ENONCE DU PROBLEME:

Le TSP est l'un des problèmes les plus étudiés dans le domaine de l'optimisation combinatoire. Le problème peut être énoncé comme suit :

Soient  $n$  villes et une matrice de distances  $C = (C_{ij})$  où  $C_{ij}$  est la distance de la ville  $i$  à  $j$ . Le problème consiste à construire le circuit Hamiltonien de longueur minimale.

Le TSP peut être vu comme un problème de la théorie des graphes en considérant les villes comme étant les sommets du graphe et les "liaisons" entre les villes comme étant les arcs du même graphe. Les poids correspondants aux distances inter-villes sont assignés à chaque arc. Mis sous cette forme, la résolution du TSP est équivalente à trouver un circuit hamiltonien de poids minimum dans ce graphe.

On distingue deux cas qui sont basés sur les propriétés de la matrice des distances (coûts)  $C$ .

\* Si  $c_{ij} = c_{ji}$  pour toute paire  $(i, j)$  appartenant à  $N = \{1, \dots, n\}$ , alors le TSP est symétrique; si non, il est asymétrique.

\* Si  $c_{ij} < c_{ik} + c_{kj}$  pour tout  $i, j$  et  $k$  appartenant à  $N$ , alors le TSP satisfait à l'inégalité triangulaire. Finalement, un TSP est euclidien si chaque noeud  $i$  de  $N$  correspondant à un point  $(x_i, y_i)$  dans le plan et  $d_{ij}$  est la distance euclidienne entre les deux points  $i$  et  $j$ . Les TSP's euclidiens sont par conséquent symétriques et satisfont à l'inégalité triangulaire.

### I.4. DEFINITION DE L'OPTIMISATION COMBINATOIRE: [20]

L'optimisation combinatoire est la branche qui étudie les problèmes définis comme suit:

Un problème d'optimisation combinatoire est défini à partir

-d'un ensemble fini  $S$ .

-d'une application  $F:S \longrightarrow R$ .

Il s'agit de déterminer  $s$  de  $S$  tel que:

$F(s)=\text{Opt}[F(x)]$ , avec  $x$  appartenant à  $S$ .

### I.5. CLASSIFICATION DU PROBLEME (au sens théorie de la complexité):

C'est durant la dernière décennie que la théorie de "la complexité des algorithmes" a été développée à partir des travaux de Cook et Karp [21]. Elle s'intéresse entre autres, à l'étude des performances des algorithmes.

#### I.5.1. Théorie de la complexité:

Cette théorie est basée sur des méthodes rigoureuses pour évaluer les algorithmes et classifier les problèmes en deux classes: difficiles et faciles.

#### I.5.2. Algorithmes et complexité:

1°/ Algorithmes : un algorithme est une procédure pas à pas pour la résolution d'un problème. Un problème peut être vu comme étant un domaine contenant des données (instances) du problème et un questionnaire sur les instances.

Par exemple, Les données du TSP sont:

-Le nombre de villes ( $n$ ).

-Une matrice des distances  $C=(C_{ij}) (n \times n)$ .

Et la question est :

Quel est le plus court circuit qui passe par toutes les villes une fois et une seule?

2°/Evaluation des algorithmes :[annexe 3].

Définition 1 :\* Etant donné deux fonctions  $f, g : \mathbb{N} \longrightarrow \mathbb{N}$ , on dit que  $f$  est  $O(g)$  s'il existe une constante  $c$  telle que : (1)  $|f(n)| \leq c \cdot |g(n)|$  pour tout  $n \in \mathbb{N}$ .

\* Une fonction  $f$  est dite "polynomiale" si elle est  $O(g)$  et si  $g$  est un polynôme en  $n$ , ou encore s'il existe deux constantes  $c$  et  $k$  telles que : (2)  $f(n) \leq c \cdot n^k$  pour tout  $n \in \mathbb{N}$ .

\* Une fonction polynomiale  $f$  est dite en " $n^\rho$ " si  $\rho$  est égal au plus petit scalaire  $k$  pour lequel (2) est satisfait.

Définition 2 : Un algorithme est dit "polynomial" si le nombre d'opérations élémentaires nécessaire pour résoudre un problème de taille  $n$  est une fonction polynomiale en  $n$ . Un algorithme est considéré comme "efficace" si, et seulement si, il est polynomial.

La notion d'efficacité est liée à l'effort de calcul (temps de calcul) nécessité par l'algorithme de résolution, lequel effort dépend de la taille du problème. Dans ce cas, le temps d'exécution sera exprimé en fonction de la taille du problème. Cependant, il existe certains algorithmes dont le temps d'exécution variera de façon significative, sans qu'il y ait variation de la taille du problème traité.

3°/Classification des problèmes:

Il existe plusieurs classes de problèmes (P, NP, NP-complet, NP-hard) dont la théorie de la complexité fait sa vocation.

En raison des nombreux détails qui découleraient fatalement du développement de ces différentes classes et qui nous

ameneraient à des considérations dépassant le stricte cadre de la présente étude, nous laisserons, au lecteur intéressé, le soin de consulter les ouvrages [13] et [20].

En ce qui concerne le TSP, Karp et Papadimitriou, cités dans [1], ont montré qu'il appartient à la classe NP-complet. Il est considéré comme étant un problème difficile conformément à la classification suivante:

- \* Les problèmes faciles, pour lesquels il existe des algorithmes de résolution efficaces.

- \* Les problèmes difficiles, dont la résolution ne semble pas admettre d'algorithmes efficaces. La résolution exacte de ce type de problème, lorsqu'elle est possible, s'effectue en un temps de calcul irraisonnable.

- \* Les problèmes "indécidables", au sens où il ne peut exister d'algorithmes pour les résoudre, qui sont évidemment difficiles.

II/APERCU SUR LES TRAVAUX DE RECHERCHE :

A cause de sa difficulté de résolution optimale (le TSP est NP-complet [1]) il a fallu attendre 1954 pour que Dantzig, Fulkerson et Johnson [12] proposent la première formulation du problème (cas symétrique). Ces chercheurs ont alors élaboré un mécanisme de résolution, dont le principe a servi de base à de nombreux autres chercheurs par la suite: celui de la résolution d'un programme linéaire en nombres entiers par relaxation de contraintes.

II.1. LES DIFFERENTES FORMULATIONS DU TSP:

Comme signalé auparavant, Dantzig et al. ont été parmi les premiers à proposer une formulation du TSP sous forme d'un problème d'affectation (1-4) avec contraintes d'intégrité (5) et contraintes d'élimination de sous-tours c'est à dire, les circuits qui ne passent pas par tous les sommets (6).

$$\text{MIN } Z = \sum_i \sum_j C_{ij} X_{ij} \quad (1)$$

Sujet à:

$$\sum_i X_{ij} = 1 \quad j=1, \dots, n \quad (2)$$

$$\sum_j X_{ij} = 1 \quad i=1, \dots, n \quad (3)$$

$$0 \leq X_{ij} \leq 1 \quad \text{pour tout } i \text{ et } j \quad (4)$$

$$X_{ij} \text{ entier} \quad \text{pour tout } i \text{ et } j \quad (5)$$

$$\sum_{i,j \in Q} X_{ij} \leq |Q| - 1 \quad \text{pour tout } Q \subseteq \{1, \dots, n\} \quad (6)$$

et  $2 \leq |Q| \leq n-1$

\* Les coefficients  $C_{ij}$  représentent le coût (ou distance) du

du voyage du noeud  $i$  vers le noeud  $j$ ; nous considérons que  $C_{ii}=0$  pour tout  $i$ .

\* Les variables  $X_{ij} = \begin{cases} 1 & \text{s'il y a une route entre } i \text{ et } j \\ 0 & \text{sinon} \end{cases}$

\* Les contraintes (2) et (3) forment les contraintes d'affectation et assurent que chaque noeud est visité une fois et une seule.

-Dantzig et al. utilisent  $O(n^2)$  variables binaires et  $O(2^n)$  contraintes avec les contraintes d'élimination de sous-tours (6).

-Miller, Tucker et Zemlin présentent une autre formulation basée sur le modèle d'affectation mais utilisant un nombre polynômial de contraintes d'élimination de sous-tours en introduisant  $O(n^2)$  variables supplémentaires continues, qui sont:

$$U_i - U_j + nX_{ij} \leq n-1 \quad \text{pour tout } j \neq i \quad (7).$$

-Gravish et Graves proposent une formulation utilisant  $O(n^2)$  variables binaires  $X_{ij}$ ,  $O(n^2)$  variables continues  $Z_{ij}$  et  $n^2 + 3n$  contraintes. Les  $Z_{ij}$  décrivent le flux (lorsqu'on a à faire à un seul type d'article) provenant des différents noeuds vers le noeud  $j$ . La contrainte d'élimination des sous-tours est:

$$\sum_{i,j} Z_{ij} - \sum Z_{ji} = 1 \quad i=2, \dots, n \quad (8)$$

$$Z_{ij} \leq (n-1)X_{ij} \quad i=2, \dots, n ; j=1, \dots, n \quad (9)$$

$$Z_{ij} \geq 0 \quad \text{pour tout } i, j \quad (10)$$

\*Formulation avec plusieurs types d'articles (produits, objets):

-Wang présente une formulation pour le TSP utilisant le flux de  $2(n-1)$  articles  $Y^k=(Y_{ij}^k)$   $k=2, \dots, n$  et  $Z^k=(Z_{ij}^k)$   $k=2, \dots, n$

$$\text{Min } \sum_i \sum_j C_{ij} \cdot X_{ij}$$

$$\sum_j \left( Y_{ij}^k - Y_{ji}^k \right) = \begin{pmatrix} 1 & \text{si } i=1 \\ -1 & \text{si } i=k \\ 0 & \text{si } i \neq k \text{ et } i \neq 1 \end{pmatrix} \quad k=2, \dots, n \quad (11)$$

$$\sum_j \left( Z_{ij}^k - Z_{ji}^k \right) = \begin{pmatrix} -1 & \text{si } i=1 \\ 1 & \text{si } i=k \\ 0 & \text{si } i \neq 1 \text{ et } k \end{pmatrix} \quad k=2, \dots, n \quad (12)$$

$$Y_{ij}^k \leq X_{ij} \quad , \quad Z_{ij}^k \leq X_{ij} \quad \text{pour tout } i, j, k \quad (13)$$

$$Y_{ij}^k \geq 0 \quad , \quad Z_{ij}^k \geq 0 \quad \text{pour tout } i, j, k \quad (14)$$

Les contraintes (11) et (12) assurent qu'une unité d'articles du type  $Y^k$  est transférée du noeud 1 (source de l'article  $Y^k$ ) vers le noeud  $k$  lorsque une unité d'article de  $Z^k$  est transférée du noeud  $k$  vers le noeud 1. Cette formulation utilise  $O(n^2)$  variables binaires,  $O(n^3)$  variables continues et  $2(n^3+n^2+n)$  contraintes.

\*Formulation avec un flux de deux types d'articles :

-Finke, Claus et Gunn proposent cette formulation pour le TSP: le voyageur de commerce quitte le noeud 1 en portant  $(n-1)$  unités d'articles de type  $Y$  et zéro article du type  $Z$ . Il distribue une unité de  $Y$  et ramasse une unité de  $Z$  à chaque noeud

de sorte que le voyageur de commerce voyage toujours avec un total de  $(n-1)$  unités. Le modèle est donné comme suit:

$$\text{MIN } \sum_i \sum_j C_{ij} (Y_{ij} + Z_{ij}) / (n-1) \quad (17)$$

$$\sum_j (Y_{ij} - Y_{ji}) = \begin{cases} n-1 & \text{pour } i=1 \\ 1 & \text{sinon} \end{cases} \quad (18)$$

$$\sum_j (Z_{ij} - Z_{ji}) = \begin{cases} -(n-1) & \text{pour } i=1 \\ 1 & \text{pour } i \neq 1 \end{cases} \quad (19)$$

$$\sum_j (Z_{ij} + Y_{ij}) = n-1 \quad \text{pour tout } i \quad (20)$$

$$Y_{ij} + Z_{ij} \in \{0, n-1\} \quad \text{pour tout } i, j \quad (21)$$

$$Y_{ij} \geq 0, \quad Z_{ij} \geq 0 \quad \text{pour tout } i, j \quad (22)$$

Les contraintes (18) et (19) garantissent la conservation de flux pour chaque type d'article. Les contraintes (20) et (21) assurent l'existence d'un seul arc qui supporte un flux composé de  $(n-1)$  unités de 2 types. La structure de l'offre et de la demande montre qu'il existe un chemin du noeud 1 vers le noeud  $j$  et un autre de  $j$  à 1. La relaxation du problème linéaire formulé précédemment est obtenue en remplaçant (21) par  $Y_{ij} + Z_{ij} \leq n-1$  pour tout  $i, j$ . Ces contraintes sont redondantes à cause de (20) et peuvent être ainsi éliminées. Le problème linéaire correspondant noté FCG (Finke, Claus et Gunn) contient  $3n$  contraintes alors que la formulation en nombres entiers contient  $O(n^2)$  contraintes en plus. Voici brièvement les plus importantes formulations du TSP. Nous entamerons, par la suite, les différentes approches de résolution. Le résumé de ces formulations se trouve en annexe 1.

## II.2. APPROCHES DE RESOLUTION:

Il existe deux approches de résolution, l'une exacte et l'autre heuristique et à cause de la difficulté de résolution optimale du TSP, les premières ont une portée pratique insuffisante. Pour cela, nous nous concentrerons beaucoup plus sur les approches heuristiques.

### II.2.1. Approches exactes :

Elles concernent:

-La méthode des plans sécants (*cutting planes*), celle de la séparation et évaluation (*branch and bound*) et leur combinaison qui donne des constructions algorithmiques impressionnantes. Cette combinaison a été utilisée par Grötschel(1972a) pour un problème à 120 villes, ensuite améliorée par Crowder et Padberg(1979b) pour atteindre une taille de 318 villes .

-Les techniques de programmation dynamique, qui s'attaquent à des problèmes de petites tailles, la méthode fut utilisée notamment par Bellman(1962) et Held & Karp(1962) .

-Relaxation lagrangienne, etc... Ces méthodes consistent à trouver des bornes inférieures utilisées ensuite dans un algorithme de *branch & bound*.

### II.2.2. Approches heuristiques :

Elles se répartissent en trois classes principales:

-Procédures de construction de tournées (*Tour construction procedures*).

-Procédures d'affinages de tournées (*Tour improvement procedures*) et

-Procédures mixtes (*Composite procedures*).

1°/Procédures de construction de tournées: (*Tour construction procedures*)

Les plus connues de ces procédures sont:

a/*Nearest neighbor algorithms*: (Rosenkrantz, Stearns & Lewis, 1977).

Etape 1: Commencer la tournée à partir d'un noeud quelconque ;  
 Etape 2: Trouver le sommet le plus proche du dernier sommet de la tournée et l'ajouter à celle-ci ;  
 Etape 3: Répéter le pas 2 jusqu'à épuisement des sommets, puis joindre le sommet initial au sommet final.

b/*Clarck and Wright savings*:

Basée sur le principe de l'économie, cette procédure agit comme suit:

Etape 1: Retenir un sommet quelconque, noté 1 ;  
 Etape 2: Calculer les économies  $S_{ij} = C_{il} + C_{lj} - C_{ij}$  pour  $i, j = 2, \dots, n$  et les classer par ordre décroissant ;  
 Etape 3: Relier les sommets selon l'ordre ainsi établi.

c/*Insertion procedures*:

Une procédure d'insertion considère une tournée formée de  $k$  sommets, à la  $k^{\text{ème}}$  itération, et détermine le prochain sommet à insérer (étape de sélection) puis l'endroit d'insertion (étape d'insertion). Il existe plusieurs types d'insertion:

-*Nearest insertion*:

1. Initialiser la tournée avec le sommet  $i$
2. Trouver le sommet  $k$  tel que  $C_{ik}$  soit minimal et former la tournée  $i \rightarrow k \rightarrow i$

d'heuristiques incluant *nearest neighbor*, *nearest insertion*, *farthest insertion* et *nearest addition* pour 12 problèmes non euclidiens, dont seulement deux satisfont à l'inégalité triangulaire. Les auteurs ont conclu que l'approche *farthest insertion method* est supérieure (meilleure) aux trois autres méthodes testées.

## 2 / Procdures d'affinage: *Tour improvment procedures*

Les meilleures heuristiques pour le TSP sont probablement celles qui consistent à échanger (permuter) les arcs dans une tournée.

Lin introduisit d'abord les procédures 2-opt et 3-opt puis définit en collaboration avec Kernighan, la procédure générale  $r$ -opt  $r \geq 3$ . Ces procédures agissent comme suit:

Etape 1: trouver une tournée initiale (généralement de manière aléatoire).

Etape 2: améliorer la solution (tournée) en substituant à un ensemble de  $r$  arcs, un autre ensemble de même nombre d'arcs.

Etape 3: répéter l'étape 2 jusqu'à ce qu'il n'y ait plus de changement.

La solution est, alors, dite  $r$ -optimale.

### \*Procédure or-optimale:

Développée par OR, cette procédure est similaire en principe à  $r$ -opt mais considère à la fois toutes les opérations d'échange portant sur deux arcs ou jonctions et celles qui conduisent à insérer une chaîne d'au plus trois sommets adjacents entre deux autres sommets du graphe. Elle est plus rapide que la première.

## 3 / Procdures mixtes : *composite procedures*

Le principe de base est le suivant:

Etape 1: Obtenir une tournée initiale, en utilisant l'une des procédures constructives.

Etape 2: Lui appliquer 2-opt.

Etape 3: Lui appliquer 3-opt.

IL existe d'autres variantes qui consistent à ne pas passer par l'une des étapes (2) ou (3), ou à réitérer ces mêmes étapes un certain nombre de fois.

Parmi elles, nous avons:

CCAO-Convex hull, Cheapest insertion, Angle selection & Or-optimale procedure.

CC-Convex hull Cheapest insertion procedure.

GA-Greatest Angle procedure.

CCA-Convex hull Cheapest insertion Angle selection procedure.

Pour illustrer ces méthodes, qui s'appliquent pour des TSP euclidiens, nous allons développer la procédure CCAO.

ALGORITHME CCAO:

Etape 1: Former la coque convexe (Convex hull) de l'ensemble des villes. Ceci donnera une sous-tournée initiale.

Etape 2: (insertion) pour toute cité  $k$  non encore incluse dans la sous-tournée initiale, identifier les deux villes (sommets) adjacentes  $i_k$  et  $j_k$  faisant partie de la sous-tournée telles que:

$Ci_k k + Ck j_k - Ci_k j_k$  soit minimale.

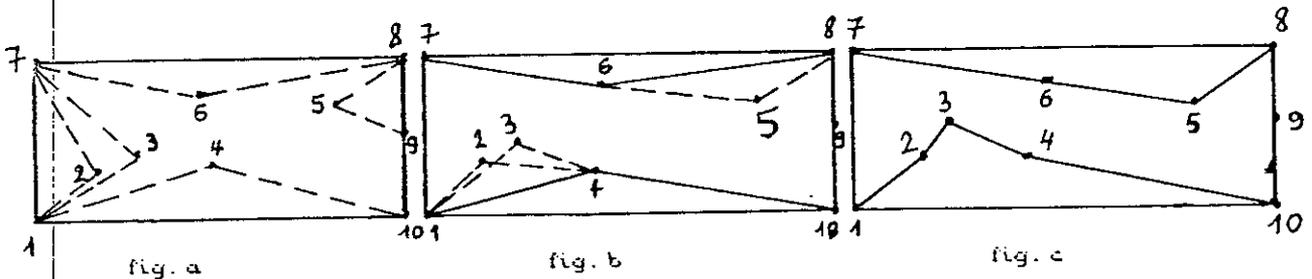
Etape 3: (selection) choisir la ville  $k^*$  qui minimise l'angle entre les deux arêtes  $\{(i_{k^*} k^*)$  et  $(k^*, j_{k^*})\}$  dans la sous-tournée et l'insérer entre  $i_{k^*}$  et  $j_{k^*}$ .

Etape 4: répéter les pas 2 et 3 jusqu'à ce que le cycle hamiltonien soit obtenu.

Étape 5:appliquer la procédure OR-opt à la tournée générée dans les étapes 1 à 4. Quand il n'y a plus d'amélioration, STOP.

Exemple: CCAO [13]

En premier, la coque convexe est générée en produisant une sous-tournée initiale constituée des villes (1,7,8,9,10).



A l'étape 2, chaque ville interne (2,3,4,5,6) est liée à une paire de villes liées à la sous-tournée déjà générée (fig.a).  
A l'étape 3, les lignes en pointillés qui donnent le plus grand angle identifient la ville à insérer (dans ce cas la ville 6).

La fig.b indique la solution intermédiaire après les deux premières insertions (villes 6 et 4 dans cet ordre). Notez que les villes 2,3 et 5 sont toutes associées à deux nouvelles paires de villes.

La fig.c montre la tournée finale après la première étape, formée des pas 1 à 4 après insertion des villes 2,5 et 3 respectivement, on applique OR-opt ensuite.

Dans l'exemple précédent, la procédure OR-opt ne donne pas d'amélioration puisque la solution est optimale dans ce cas.

### II.2.3. Autres méthodes:

Nous citerons deux de ces méthodes.

1° Méthode de recherche avec tabous: (G.Laporte, 1989, Montréal

University).

On examine successivement les solutions voisines d'une solution donnée et on permet la détérioration de l'objectif pour éviter que le processus ne se bloque dans un optimum local. Afin d'empêcher le processus de cycler, on interdit les solutions déjà examinées que l'on place dans une liste tabous constamment mise à jour.

2° Méthode de "recuit simulé" : (B.L.Golden et S.Skiscim, 1986).

Cette méthode est due à des spécialistes de physique statistique qui ont transposé à la résolution des problèmes d'optimisation combinatoire les techniques de "recuit". Schématiquement, le principe est le suivant: pour amener un matériau à un état solide d'énergie minimum, on le chauffe suffisamment jusqu'à ce que les particules soient distribuées aléatoirement dans l'état liquide.

Puis, afin d'éviter des minima locaux, on en abaisse graduellement la température par palier jusqu'à ce que le système atteigne son état d'équilibre à l'intérieur d'un palier donné.

A haute température, on peut atteindre tous les états mais, à mesure que le système refroidit le nombre de possibilités diminue et le processus finit par converger vers un état stable.

En optimisation combinatoire, on cherchera à passer d'une solution initiale donnée à une solution de coût minimum, en effectuant des modifications graduelles. L'efficacité de la méthode est due au fait que chaque calcul élémentaire est relativement simple. On sait que si la décroissance de la température est bien contrôlée, on obtient des résultats spectaculaires. Ainsi Kirk Patick & Alii ont résolu par ce type d'approche un TSP avec 3200 villes!

II.3. RAPPELS SUR LA THEORIE DES GRAPHES:a/ Définition d'un graphe et champ d'application:

\* Un graphe  $G=[X,U]$  est déterminé par la donnée:

-D'un ensemble  $X$  dont les éléments sont appelés des sommets ou des noeuds. Si  $N=|X|$  est le nombre de sommets, on dit que le graphe  $D$  est d'ordre  $N$ . Les sommets sont numérotés de 1 à  $N$ .

-D'un ensemble  $U$  dont les éléments  $u \in U$  sont des couples ordonnés de sommets appelés arcs. Si  $U=(i,j)$  est un arc de  $G$ ,  $i$  est l'extrémité initiale et  $j$  l'extrémité terminale de  $U$ . On notera souvent  $|U|=M$ .

\* Le langage des graphes permet de représenter simplement la structure d'un grand nombre de situations:

-L'exemple le plus classique est la représentation d'un réseau de communications: réseaux de routes, représentés par une carte routière, réseaux de chemin de fer, de téléphone, de relais de TV, réseaux électriques, réseaux des informations dans une organisation, etc...

-La famille d'exemples la plus générale est la représentation d'une relation binaire :algébrique, mécanique, chimique, sociologique, etc...comme les règles de certains flux(dames ou échecs), la supériorité des participants à un tournoi, l'ordonnancement des opérations de montage ou de démontage d'un ensemble technologique, etc...

-un cas particulier important est la représentation d'une évolution, comme le passage d'un certain état à un autre, les chaînes de Markov, ou encore l'évolution d'une population dans un phénomène démographique.

b/Définition des chemins et circuits hamiltoniens:

\* Un circuit hamiltonien (resp. cycle hamiltonien) est un circuit (resp. un cycle) qui passe une fois, et une fois seulement, par chacun des sommets d'un graphe  $G[X, U]$ .

\* L'étude des problèmes hamiltoniens (recherche de circuits ou cycles hamiltoniens) remonte aux origines de la théorie des graphes.

L'intérêt porté aujourd'hui à ces problèmes s'explique par leurs applications nombreuses: tournées de distribution (TSP et ses extensions), tracé automatique sur ordinateur, problèmes d'ordonnancement d'ateliers, etc...

\* Bien que ces problèmes apparaissent très semblables par leur formulation, ils se révèlent fort différents du point de vue de leur résolution.

\* Pour  $n+1$  sommets, il y a  $n!$  circuits hamiltoniens possibles, donc  $n!$  coûts de circuits à calculer, il "suffit" alors de retenir le minimum de ces coûts et le problème est ainsi résolu. Mais, au delà d'une certaine valeur de  $n$ , ( $n=20$  par exemple) les ordinateurs les plus puissants sont incapables de traiter ces problèmes en un temps raisonnable. Nous donnons un exemple qui fera sentir l'incapacité de le résoudre avec l'algorithme déjà cité, en le comparant à un algorithme trivial pour résoudre le problème de la recherche du plus petit nombre stocké dans un fichier de  $N$  nombres. Cet algorithme consiste à comparer le premier et le second élément, comparer le minimum

obtenu au troisième et ainsi de suite ; on effectue ainsi (N-1) comparaisons (on montre que cet algorithme est le meilleur au sens où il n'en existe pas d'autre qui permette de résoudre le problème en moins d'opérations). La différence de performances apparaît très clairement [tableau 1] (l'ordinateur utilisé est supposé pouvoir effectuer un milliard d'opérations par seconde).

n=nombre de villes	10	20	30	40	50	60
$n^2$ =nombre de données	100	400	900	1600	2500	3600
Durée d'exécution de l'algorithme permettant d'extraire le min de $N=n^2$ nombres(en $\mu s$ )	0.1	0.4	0.9	1.6	2.5	3.6
Durée d'exécution de l'algorithme permettant d'extraire le min de $N=(n-1)!$ nombres (en années).	$10^{-8}$	$10^0$	$10^{16}$	$10^{30}$	$10^{45}$	$10^{57}$

tableau 1: Tableau comparatif des deux algorithmes [20]

Notons que cette différence de comportement peut être rattachée au fait que le TSP a une structure combinatoire qui fait qu'en spécifiant les  $n^2-n$  distances intervilles, on engendre (de manière implicite, évidemment, un nombre sans commune mesure de solutions possibles.

#### II. 4. CONCLUSION :

Même si les méthodes heuristiques constituent un palliatif

efficace pour la taille importante du problème, elles véhiculent, cependant, des difficultés qui se résument ainsi: -On demande en général à une heuristique de conduire "rapidement" (c-a-d efficacement) à une "bonne solution". Ce dernier terme nous attire la première difficulté, en effet, comment dire des choses précises sur des êtres par nature approximatifs?.

-Les méthodes approximatives sont très souvent spécifiques à des problèmes (cas particulier du TSP) qu'elles se proposent de résoudre.

-L'évaluation des performances de ces heuristiques (qualité de la solution qui pourrait en sortir, temps de calcul...).

Dans le but de surmonter la dernière difficulté, les chercheurs ont énormément travaillé sur ce point moyennant plusieurs outils. Parmi ces derniers, on trouve ce qu'ils nomment l'optimisation statistique [15].

Il est à signaler en dernier lieu, qu'il n'existe pas, à notre connaissance, d'algorithme exact et efficace pour la résolution du TSP.

III/L'APPROCHE STATISTIQUE DU PROBLEME DU VOYAGEUR DE COMMERCE.III.1. INTRODUCTION:

L'un des mystères, entourant le TSP, est la remarquable performance des heuristiques simples. Les propriétés de ces dernières sont usuellement établies par des méthodes empiriques, qui consistent à trier les heuristiques et observer la qualité des résultats fournis par celles-ci.

Parmi les méthodes empiriques utilisées dans ce domaine, nous allons focaliser notre attention sur les différentes méthodes statistiques, particulièrement, celles qui consistent à estimer un intervalle de confiance de la solution optimale. Un aperçu sur différentes tentatives de construction de l'intervalle sera donné; puis une méthode utilisant les moments sera développée.

III.2. HISTORIQUE:

Le nombre d'articles publiés dans le domaine de l'analyse empirique des heuristiques est énorme. Les auteurs ont utilisé différents tests (test de Wilcoxon, test de Friedman) pour comparer la performance des heuristiques, sans arriver à un résultat concluant. Ils se sont penchés sur d'autres méthodes telles que: Approche de l'utilité esperée (expected utility approach) [13]. Ils ont aussi travaillé sur l'estimation statistique de la solution optimale d'un assez grand nombre de problèmes rentrant dans le domaine de l'optimisation combinatoire. Pour asseoir leurs méthodes la plupart d'entre eux, ont eu recours au classique problème du voyageur de commerce.

En ce qui concerne l'approche de l'utilité espérée, les heuristiques utilisées sont de trois types: *Tour construction procedure, tour improvement procedure, composite procedure*).

GOLDEN et al. ont abouti aux conclusions suivantes:

1-Plusieurs procédures de construction de tournées peuvent trouver une solution avec 5 % à 7% de l'optimalité. Ces procédures sont utilisées quand une solution effective est désirée.

2-Les procédures d'affinage, en particulier 2-opt et 3-opt avec des solutions initiales aléatoires, opèrent avec approximativement la même efficacité et efficacité que la meilleure des procédures de construction.

3-La procédure mixte peut trouver une solution du TSP avec 2% à 3% de l'optimalité avec une grande régularité. Cependant, cette procédure de calcul demande un temps de calcul sensiblement élevé, comparé aux procédures d'affinage et de construction.

4-Pour trouver efficacement une solution du TSP avec 1% à 2% de l'optimalité, il faut répéter l'application de "three steps" (procédure mixte).

Parmi les différentes procédures d'estimation de l'intervalle de confiance on trouve: [8]

-Procédure donnée par Clough. Il a appliqué la théorie des valeurs extrêmes au problème d'estimation de l'intervalle de confiance de l'optimum global d'une fonction à plusieurs variables assujettie à certaines contraintes, en utilisant la méthode de Monte Carlo. L'intervalle est donné par la formule suivante:

$$P \left\{ \left[ \frac{S+H}{S-1} \right]^v - \left[ \frac{1+H}{S-1} \right] \left[ \frac{1}{S} \right] \sum_{i=1}^S X_i \leq a \leq v \right\} = 1 - \left[ \frac{S-1}{S-H} \right]^{S-1}$$

S: taille de l'échantillon

H: réel quelconque

a: solution optimale

$X_i$ : solutions données par la l'heuristique 2-Opt

-Une procédure, donnée par Mann et al, utilise l'approximation par la loi de Béta. L'intervalle de confiance construit dépend de l'approximation.

-Une procédure donnée par Robson et Whitlock. Elle procède comme suit: Soit  $X_i$  la  $i^{\text{ème}}$  plus petite valeur. On pose  $v=X_1$ . L'intervalle de confiance à un niveau de  $100(1-\alpha)\%$  est obtenu par une troncature de la valeur inférieure d'une distribution uniforme. Ils ont montré que cette formule donne des intervalles de confiance approximatifs pour d'autres lois de probabilité. Cet intervalle est donné par:

$$P \left[ v - \left( \frac{1-\alpha}{\alpha} \right) (x_2 - v) \leq a \leq v \right] \approx 1-\alpha$$

$\alpha$ : risque de se tromper

Ces procédures présentent des inconvénients et paraissent inefficaces. Le lecteur intéressé est invité à consulter l'ouvrage [8].

Une procédure d'estimation d'un intervalle de confiance semble efficace. Elle fait appel à un théorème suivant:

"La distribution du minimum d'un échantillon de taille  $m$  tend vers la loi de Weibull lorsque la taille de l'échantillon tend vers l'infini" [5].

Cette procédure sera développée dans ce qui suit et utilisée pour comparer les résultats qu'elle donne avec ceux obtenus par la méthode des moments.

### III.3. ESTIMATION DE LA SOLUTION OPTIMALE DU TSP PAR INTERVALLE DE CONFIANCE:

#### 1°/Enoncé du problème:

Habituellement, toutes les procédures utilisent des approches approximatives pour résoudre le TSP. Nous allons nous intéresser à un domaine nommé l'optimisation statistique [15].

C'est une approche qui consiste, non pas à trouver une solution approchée des problèmes combinatoires, mais à estimer la valeur optimale en utilisant une loi de probabilité.

Dans notre travail, nous nous intéresserons au TSP moyennant l'utilisation de la loi de Weibull et la "méthode des moments".

#### 2°/Définition d/e la loi de Weibull:

\* Définition: une v.a X est distribuée selon une loi de Weibull de paramètres :  $a \geq 0$  ;  $b > 0$  ;  $c \geq 0$  si et seulement si :

- Elle est continue.
- Elle prend des valeurs dans l'intervalle  $[a, \infty[$ .
- Sa densité de probabilité est donnée par :

$$f(x) = \frac{c}{b} \left[ \frac{x-a}{b} \right]^{c-1} \exp \left[ - \left[ \frac{x-a}{b} \right]^c \right]$$

avec : a : paramètre de position  
 b : paramètre d'échelle  
 c : paramètre de forme

Caractéristiques:

$$\text{-Moyenne} = E[X] = a + \frac{b}{c} \Gamma \left( \frac{1}{c} \right) = a + b \Gamma \left( 1 + \frac{1}{c} \right)$$

$$\text{-Var}[X] = E[(X - M_x)^2] = b^2 \Gamma \left( 1 + \frac{2}{c} \right) - b^2 \left[ \Gamma \left( 1 + \frac{1}{c} \right) \right]^2$$

$$\text{-}E[X^3] = b \Gamma \left( \frac{3}{c} + 1 \right) + 3b^2 a \Gamma \left( \frac{2}{c} + 1 \right) + 3b a^2 \Gamma \left( \frac{1}{c} + 1 \right) + a^3 \Gamma(1)$$

$$\Gamma(x) = \int_0^{\infty} \exp(-u) u^{x-1} du$$

3°/Procédure d'estimation:

En parcourant les travaux de recherche réalisés durant ces quarante dernières années, on constate que presque tous les domaines susceptibles d'apporter une solution au problème du TSP ont été explorés de manière systématique. Parmi ces domaines, on trouve ce que l'on nomme "l'appréciation statistique des

heuristiques" ou "analyse empirique des heuristiques" (*empirical analysis of the heuristics*) dont l'un d'eux est l'estimation statistique de la longueur optimale d'un circuit.

L'estimation statistique de la longueur optimale d'un circuit:

Initialement étudiée par KLEIN(1975), l'estimation statistique de la longueur optimale des circuits a ensuite été traitée par de nombreux auteurs tel que: GOLDEN(1977b), DANNENBRING(1977), GOLDEN(1978), GOLDEN & ALT(1979), ARIYAWANSA(1980), LOS & LARDINOIS(1982), DERIGS(1983) [13]. Ces différents chercheurs ont fait appel à un résultat, trouvé par FISHER & TIPETT(1928), fondamental dans le cadre de la théorie de distribution statistique des valeurs extrêmes, afin de déterminer plus exactement une valeur estimée de l'optimum [Mac ROBERTS(1979), GOLDEN(1977b, 1978)]. Le résultat fondamental trouvé par FISHER & TIPETT est:

Supposons que l'on dispose de  $S$  échantillons indépendants (tirés de façon indépendante), tous de même taille  $m$ , issus d'une population continue et bornée inférieurement par  $a$ .

Si  $X_i$  est la plus petite valeur d'un échantillon  $i$  alors soit:  $v = \text{MIN} \{ X_i / 1 \leq i \leq S \}$ .

Lorsque  $m$  devient grand, la distribution de  $X_i$  est proche de la distribution de Weibull avec  $a$ : paramètre de position. La fonction de répartition est:

$$F_x(x_0) = P\left[x \leq x_0\right] = 1 - \exp\left[-\left(\frac{x-a}{b}\right)^c\right]$$

$$x_0 \geq a > a ; b > 0 ; c \geq 0.$$

b: paramètre d'échelle.

c: paramètre de forme.

La population mère est composée de  $n!$  circuits avec un coût total non inférieur à  $a$ , qui est la longueur optimale du circuit

La distribution est discrète, mais on peut approximer son histogramme par une distribution continue lorsque le nombre de tournées (circuits) est grand [15].

Ces chercheurs ont ensuite vérifié l'hypothèse que  $X_i$  est distribuée selon une loi de Weibull, avec un risque de  $\alpha=0.05$ , en utilisant la statistique de KOLMOGOROV-SMIRNOV [GOLDEN(1978)] sous l'hypothèse d'indépendance.

\*L'estimation ponctuelle des paramètres de la loi de Weibull:

L'estimation ponctuelle des paramètres d'une loi de probabilité peut se faire de deux manières:

a/Méthode des moments.

b/Méthode de maximum de vraisemblance.

-Les moments de la loi de Weibull sont donnés par des formules très compliquées. C'est pour cela que la méthode de maximum de vraisemblance est préférée.

-La fonction de maximum de vraisemblance est donnée par:

$$L(x_1, x_2, \dots, x_n; a, b, c) = \prod_{i=1}^n f(x_i)$$

$$= \prod_{i=1}^n \left[ \frac{c}{b} \right] \left[ \frac{x_i - a}{b} \right]^{c-1} \exp \left[ - \left( \frac{x_i - a}{b} \right)^c \right]$$

$$\text{Log}L = n \text{Log}c + nc \text{Log}b + (c-1) \sum_{i=1}^n \text{Log}(x_i - a) - \sum_{i=1}^n \left[ \frac{x_i - a}{b} \right]^c$$

En dérivant LogL par rapport à  $a$ ,  $b$  et  $c$ , et en égalisant à zéro, on obtient les équations du maximum de vraisemblance:

$$\frac{\partial \text{LogL}}{\partial a} = -(c-1) \sum_{i=1}^n \left[ \frac{1}{X_i - a} \right] + \frac{c}{b^c} \sum_{i=1}^n (X_i - a)^{c-1} = 0 \quad (1)$$

$$\frac{\partial \text{LogL}}{\partial b} = -\frac{Nc}{b} + \frac{c}{b^{c+1}} \sum_{i=1}^n (X_i - a)^c = 0 \quad (2)$$

$$\frac{\partial \text{LogL}}{\partial c} = \frac{n}{c} - n \text{Log} b + \sum_{i=1}^n \text{Log}(X_i - a) - \sum_{i=1}^n \left[ \frac{X_i - a}{b} \right]^c \text{Log} \left[ \frac{X_i - a}{b} \right] = 0 \quad (3)$$

De l'équation (2) on obtient:

$$b = \left( \sum_{i=1}^n \frac{X_i - a}{b} \right)^{1/c} \quad (*)$$

En substituant (\*) dans (1) et (3) on obtient:

$$-(c-1) \sum_{i=1}^n \left[ \frac{1}{X_i - a} \right] + nc \sum_{i=1}^n (X_i - a)^{c-1} / \sum_{i=1}^n (X_i - a)^c = 0 \quad (4)$$

$$\frac{n}{c} + \sum_{i=1}^n \text{Log}(X_i - a) - n \sum_{i=1}^n (X_i - a)^c \text{Log}(X_i - a) / \sum_{i=1}^n (X_i - a)^c = 0 \quad (5)$$

Une approche consiste à résoudre les deux équations (4) et (5) par des méthodes numériques, ainsi on aura des estimations de  $a$  et  $c$  puis en remplaçant  $a$  et  $c$  par leurs valeurs dans (\*) on aura une estimation de  $b$

Une seconde approche consiste à maximiser  $\text{LogL}$  sous les contraintes  $b > 0$  et  $c > 0$ , pour cela, différentes méthodes peuvent être utilisées:

-Méthode des barrières.

-Méthode utilisant les pénalités.

Une troisième méthode consiste à utiliser la méthode des moindres carrés (*less Combersome third line of attack based on the least squares concept*).

La fonction de répartition de la loi de Weibull donne la première équation:

$$\exp \left[ - \left( \frac{x_0 - a}{b} \right)^c \right] = 1 - F(x_0)$$

$$c \log(x_0 - a) - c \log b = \log \left[ -\log(1 - F(x_0)) \right]$$

L'équation (6) est linéaire avec  $\log(x_0 - a)$  comme variable exogène et  $\log[-\log(1 - F(x_0))]$  comme variable endogène (dépendante).

On pose:

$$\alpha = c, \quad \beta = c \log b, \quad X = \log(x_0 - a), \quad Y = \log \left[ -\log(1 - F(x_0)) \right]$$

$$(6) \text{ devient: } Y = \alpha X + \beta \quad (7)$$

Si on fixe le paramètre  $a$ , les paramètres  $\alpha$  et  $\beta$  peuvent être estimés par la méthode des moindres carrés ordinaires (MCO).

La droite (7) peut être améliorée en cherchant l'ensemble des valeurs de  $a$  qui permettent d'avoir le plus grand coefficient de corrélation linéaire.

Cet ensemble est borné puisque  $\alpha = \min \{ X_i / 1 \leq i \leq N \}$

$F(X_0)$  est estimé à partir des données.

Si le plus grand coefficient de corrélation est loin de 1 en valeur absolue, on doit vérifier si l'hypothèse qui dit que les solutions données par les heuristiques sont distribuées selon la loi de Weibull.

Remarque:

-L'approche du MCO donne des estimations des paramètres qui vérifient toutes les équations du maximum de vraisemblance avec une bonne précision.

-Le coefficient de corrélation est très proche de l'unité.

-Les paramètres étant estimés, le test de KOLMOGOROV-SMIRNOV peut être utilisé pour tester l'hypothèse nulle.

Une estimation par intervalle de confiance de  $a$ , solution du TSP, est réalisée.

La procédure de construction de l'intervalle de confiance est donnée, en s'appuyant sur la fonction de distribution de Weibull qui nous permettra d'écrire:

$$\begin{aligned} \Pr\{v \leq a+b\} &= 1 - \Pr\{v > a+b\} \\ &= 1 - \{1 - F_{x_1}(a+b)\} \{1 - F_{x_2}(a+b)\} \dots \{1 - F_{x_s}(a+b)\} \\ &= 1 - \exp(-s) \end{aligned}$$

ou

$$\Pr\{v-b \leq a \leq v\} = 1 - \exp(-s)$$

la procédure est:

-Etape 0: Trouver les solutions  $X_1, X_2, \dots, X_s$  données par l'heuristique.

-Etape 1: Réarranger les solutions par ordre croissant pour

obtenir  $v = X_1, X_2, \dots, X_s$ .

-Etape 2: Déterminer une bonne estimation initiale des paramètres  $a$ ,  $b$  et  $c$ .

-Etape 3: Résoudre les équations du maximum de vraisemblance pour obtenir les paramètres  $\hat{a}$ ,  $\hat{b}$  et  $\hat{c}$ .

-Etape 4: L'intervalle de confiance est donné par:  
 $[v - \hat{b}, v]$ .

Dans cette procédure, les tournées initiales sont générées aléatoirement par la procédure ranper [NIJENHUIS & WILF, 1975], l'heuristique d'affinage utilisée est *2-opt procedure*.

Pour l'ensemble des problèmes ainsi générés, la solution "optimale", donnée par l'heuristique, se trouve toujours dans l'intervalle de confiance  $[v - \hat{b}, v]$ . Des résultats sont donnés en annexe 2.

Comme il existe peu de TSP dont la solution optimale est connue pour une taille de problème dépassant  $n=100$ , certains auteurs proposent une formule qui donne la longueur asymptotique d'un circuit minimal lorsque les villes sont distribuées uniformément sur une aire rectangulaire décomposée en  $R$  unités. Cette formule est donnée par:

$$L(n, R) = K\sqrt{nR}$$

avec

$K$ : constante.

$n$ : nombre de villes (sommets).

$R$ : nombre d'unités de découpage.

Des bornes empiriques de la constante  $K$  sont données par [STEIN, 1977].

$$0.765 \leq K \leq 0.765 + \frac{4}{n}$$

Afin de s'assurer de l'efficacité de cette méthode, d'autres expériences de résolutions (grand nombre de solutions: on pourrait

augmenter la taille de l'échantillon  $S$  ont été effectuées.

Remarque:

-Un grand nombre d'applications numériques sur des problèmes composés du TSP et du QAP (Quadratic Assignment Problem) sont décrits par DERIGS, 1983. Ce travail confirme aussi la validité de la procédure de l'intervalle de confiance.

-L'intervalle de confiance permet de comparer la qualité de la solution des différentes heuristiques. Le rapport  $b/v$ , représentant la largeur relative de l'intervalle, nous donne une idée sur l'efficacité de l'heuristique utilisée. En effet, plus  $b/v$  est petit plus l'heuristique utilisée est puissante.

LOS & LARDINOIS, 1982 cités dans [13], ont proposé une procédure d'amélioration de l'intervalle de confiance donné par la loi de Weibull en posant:

$$[v-b/T;v] \text{ au lieu de } [v-b;v] \text{ avec } T = [-S/\text{Log}\alpha]^{1/c}$$

- $\alpha$ :risque de première espèce (de se tromper).

- $S$ :taille de l'échantillon.

- $c$ :paramètre de forme de la loi de Weibull.

Ces mêmes auteurs ont testé l'efficacité de cette méthode.

#### III.4. DETERMINATION DE LA MOYENNE ET DE LA VARIANCE DE LA LOI DE DISTRIBUTION DE LA POPULATION:

L'une des méthodes théoriques, servant à déterminer la loi de distribution d'un phénomène, consiste à déterminer, à partir d'observations, les moments de cette loi.

nous avons essayé d'appliquer cette méthode au TSP. Il apparaîtra, dans ce qui suit, que cette approche n'est pas évidente eu égard à la complexité des formules trouvées pour les moments d'ordre 1 (moyenne) et 2 (variance) seulement.

Nous commencerons notre travail en déterminant, les moments d'abord, ensuite nous donnerons les résultats, concernant la borne inférieure.

a/Calcul des moments:

-Moment d'ordre 1 (moyenne):

Notre système est représenté par le graphe suivant:  $G(X,U,T)$

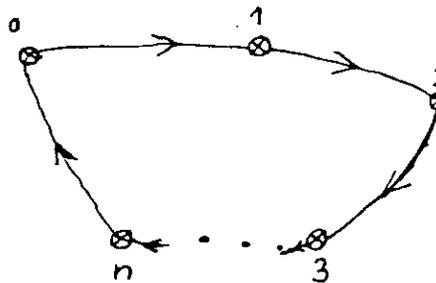


Fig.1: graphe d'un circuit Hamiltonien quelconque.

Ce graphe peut être, à son tour, représenté par la matrice des coûts  $T=(t_{ij}) \ i,j=0, \dots, n$

$$T = \begin{matrix} & \begin{matrix} 0 & \dots & 1 & \dots & n \end{matrix} \\ \begin{matrix} 0 \\ t_{10} \\ t_{n0} \end{matrix} & \begin{matrix} \dots \\ \dots \\ \dots \end{matrix} & \begin{matrix} 1 \\ 0 \\ t_{n1} \end{matrix} & \begin{matrix} \dots \\ \dots \\ \dots \end{matrix} & \begin{matrix} n \\ t_{0n} \\ 0 \end{matrix} & \begin{matrix} 0 \\ 1 \\ \vdots \\ n \end{matrix} \end{matrix}$$

Pour avoir un circuit hamiltonien quelconque, en commençant par le sommet 0, il suffit de choisir un ordre quelconque de passage par les sommets une fois et une seule.

Etant donné la configuration:  $\{0 \rightarrow 1 \rightarrow 2 \dots \rightarrow n \rightarrow 0\}$  pour obtenir tous les autres circuits, il suffit d'envisager toutes les permutations possibles.

Une méthode pour la détermination d'un circuit hamiltonien est:

1-Prendre un élément quelconque sur la ligne '0' de la matrice 'T'.

2-Barrer la ligne et la colonne de cet élément et retenir le coût  $t_{0j}$  de l'arc  $(0 \rightarrow j)$ .

3-Choisir un élément  $t_{jk}$  non barré de la ligne  $j$ , et retenir son coût  $t_{jk}$ , barrer la ligne  $j$  et la colonne  $k$ .

4-Affecter la valeur de  $k$  à  $j$ .

5-Refaire les étapes 3 et 4 jusqu'à ce que toutes les lignes et les colonnes soient barrées.

Le calcul de la moyenne se fera comme suit:

-A chaque circuit hamiltonien  $K$  est associé une valeur  $X_k$  telle que:

$X_k$  = somme de toutes les valeurs associées aux arcs du circuit  $k$ .  
La valeur moyenne de tous les circuits hamiltoniens est:

$$T_{moy} = \frac{\text{somme des valeurs de tous les circuits}}{\text{nombre de circuits hamiltoniens}} = \left\{ \sum_{k=1}^{n!} X_k \right\} / n!$$

\* Quel est le nombre de circuits hamiltoniens contenant l'arc  $(i,j)$ ? ceci revient à chercher combien de fois le sommet  $j$  vient juste après  $i$ . Puisqu'entre  $i$  et  $j$  aucun sommet n'apparaît, donc on peut représenter la séquence  $i \rightarrow j$  par un seul sommet  $k$  et (a) devient:

$$\left\{ \rightarrow 1 \rightarrow 2 \rightarrow \dots \rightarrow k \rightarrow \dots \rightarrow (n-1) \rightarrow n \rightarrow 0 \right\}$$

Il y a  $(n-1)$  sommets.

Le nombre des permutations possibles s'élève alors à  $(n-1)!$  permutations, chaque permutation donnant lieu à un circuit hamiltonien. Ainsi, le nombre de circuits hamiltoniens est  $(n-1)!$ , et par conséquent l'arc  $(i,j)$  apparaît  $(n-1)!$  sur l'ensemble de tous les circuits hamiltoniens. Ce raisonnement est valable pour tous les arcs.

\* Maintenant, calculons la somme  $\sum X_k$ , on trouve le coût de tous les arcs du graphe, qui sont répétés  $(n-1)!$  fois. Dans chaque circuit hamiltonien de la population mère formée par les  $n!$  circuits hamiltoniens, un arc quelconque  $(i,j)$  a un coût égal à  $t_{ij}$ , son coût sur l'ensemble des circuits hamiltoniens est, par conséquent, égal à  $(n-1)! * t_{ij}$ . En sommant sur tous les circuits

\* Pour calculer  $T_{moy}$  (c-a-d la moyenne), on procède de la manière suivante:

\* On calcule la somme des valeurs de tous les circuits:

\* Un circuit hamiltonien (ch) dans un graphe est défini par la donnée de l'ordre de passage par tous les sommets du même graphe. Pour avoir tous les circuits, il suffit de changer l'ordre de passage par les sommets.

Exemple: graphe à 7 sommets

Soit le circuit hamiltonien suivant 1

$\{0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6 \rightarrow 0\}$

Le sommet 0 étant le point de départ du circuit 1,

les autres circuits sont:

$\{0 \rightarrow 2 \rightarrow 1 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6 \rightarrow 0\} : 2^{\circ}$

$\vdots$

$\{0 \rightarrow 6 \rightarrow 5 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow 1 \rightarrow 0\} : 720^{\circ}$  circuit (6!).

\* Sur l'ensemble de tous les circuits hamiltoniens combien de fois apparaît un arc quelconque  $(i,j)$ ?  $i$  et  $j$  étant des sommets du graphe. Un arc  $(i,j)$  apparaît une et une seule fois dans un circuit hamiltonien, donc il suffit de connaître le nombre total de circuits hamiltoniens qui comprennent l'arc  $(i,j)$ .

Soit  $G(X,U,T)$  un graphe où:

-  $|X| = n+1$  sommets, numérotés de 0 à  $n$ .

-  $U$ : l'ensemble de tous les arcs.

-  $T$ : matrice des coûts.

Puisqu'un circuit hamiltonien passe par tous les sommets, prenons 0 comme point de départ. Soit le circuit suivant:

$\{0 \rightarrow 1 \rightarrow 2 \rightarrow \dots \rightarrow i \rightarrow j \rightarrow \dots \rightarrow (n-1) \rightarrow n \rightarrow 0\}$  (a)

hamiltoniens nous obtenons:

$$(n-1)!t_{01} + (n-1)!t_{12} + \dots + (n-1)!t_{n-1n} + (n-1)!t_{n0} = \sum_{k=1}^{n!} X_k$$

Ou encore,

$$\sum_{k=1}^{n!} X_k = (n-1)! \sum_{i=1}^n \sum_{j=1}^n t_{ij} \quad i \neq j$$

\* Comme le nombre total de circuits hamiltoniens est égal à  $n!$ , la moyenne est:

$$T_{moy} = \frac{\text{Somme des valeurs (coûts) de tous les circuits hamiltoniens}}{\text{Nombre total de circuits hamiltoniens dans le graphe}}$$

$$\frac{\sum_{k=1}^{n!} X_k}{n!} \quad (*)$$

$$T_{moy} = \bar{T} = \frac{\sum_{k=1}^{n!} X_k}{n!} = \frac{(n-1)! \sum_{k=1}^{n!} t_{ij}}{n!} \rightarrow T = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n t_{ij}$$

$$\bar{T} = \sum_{i=1}^n \sum_{j=1}^n t_{ij} \quad i \neq j$$

Donc la valeur moyenne de l'ensemble des  $n!$  circuits est égale à

Donc la valeur moyenne de l'ensemble des  $n!$  circuits est égale à la moyenne des coûts (poids) de l'ensemble des arcs.

Cette formule réduit énormément le temps de calcul de la moyenne. Elle ne fait intervenir que les coûts des arcs, alors que (\*) fait intervenir  $n!$  circuits chacun comptant  $(n+1)$  arcs.

-Calcul du moment d'ordre 2 (variance):

Posons  $X_k$  le coût total du circuit hamiltonien  $k$ .

La variance échantillonnée est donnée par la formule suivante:

$$s^2 = \frac{1}{n!-1} \sum_{k=1}^{n!} (X_k - \bar{T})^2 \quad \text{qu'on peut écrire}$$

$$s^2 = \frac{1}{n!} \sum_{k=1}^{n!} (X_k - \bar{T})^2 = m_2 - \bar{T}^2 \quad m_2 \text{ étant le moment d'ordre 2 par rapport à l'origine.}$$

Calculons  $m_2$ , pour ce faire, prenons un exemple

Soit un graphe  $G[X, U, T]; |X|=6$  ;  $X = \{ 0, 1, 2, 3, 4, 5 \}$ ;

Le nombre de circuits hamiltoniens est égal à  $5! = 120$ .

Ces circuits sont:

0123450	0213450	0312450	0412350	0512340
0123540	0213540	0312540	0412530	0512430
0124350	0214350	0314250	0413250	0513240
0124530	0214530	0314520	0413520	0513420
0125340	0215340	0315240	0415230	0514230
0125430	0215430	0315420	0415320	0514320
0132450	0231450	0321450	0421350	0521340
0132540	0231540	0321540	0421530	0521430
0134250	0234150	0324150	0423150	0523140
0134520	0234510	0324510	0423510	0523410
0135240	0235140	0325140	0425130	0524130
0135420	0235410	0325410	0425310	0524310
0142350	0241350	0341250	0431250	0531240
0142530	0241530	0341520	0431520	0531420
0143250	0243150	0342150	0432150	0532140
0143520	0243510	0342510	0432510	0532410
0145230	0245130	0345120	0435120	0534120
0145320	0245310	0345210	0435210	0534210
0152340	0251340	0351240	0451230	0541230
0152430	0251430	0351420	0451320	0541320
0153240	0253140	0352140	0452130	0542130
0153420	0253410	0352410	0452310	0542310
0154230	0254130	0354120	0453120	0543120
0154320	0254310	0354210	0453210	0543210

$$m_2 = \frac{1}{n!} \sum_{k=1}^{n!} X_k^2 = \frac{1}{n!} (X_1^2 + X_2^2 + \dots + X_{n!}^2)$$

Pour l'exemple, on a :

$$m_2 = \frac{1}{5!} \sum_{k=1}^{n!} X_k = t_{01}^2 + t_{12}^2 + t_{23}^2 + t_{34}^2 + t_{45}^2 + t_{50}^2 + 2(t_{01}t_{12} + t_{01}t_{23} + t_{01}t_{34} + t_{01}t_{45} + t_{01}t_{50} + t_{12}t_{23} + t_{12}t_{34} + t_{12}t_{45} + t_{12}t_{50} + t_{23}t_{34} + t_{23}t_{45} + t_{23}t_{50} + t_{34}t_{45} + t_{34}t_{50} + t_{45}t_{50})$$

Le développement de cette somme est très long, par conséquent, nous donnerons la formule générale de m :

Nous avons constaté que, pour un graphe à (n+1) sommets, le terme  $t_{ij}^2$  apparaît dans la somme (n-1)! fois et le terme  $t_{ij}t_{hl}$  apparaît 2(n-2)! fois

Donc

$$s^2 = \frac{1}{n} \sum_{i \neq j} t_{ij}^2 - \frac{1}{n^2} \sum_{i \neq j} t_{ij}^2 + \frac{2}{n(n-1)} \sum \sum t_{ij} t_{jh} - \frac{1}{n^2} \left[ \sum \sum t_{ij} t_{hl} + \sum \sum t_{ij} t_{hj} + \sum \sum t_{ij} t_{ih} \right]$$

$$s^2 = \frac{n-1}{n^2} \sum t_{ij}^2 + \frac{n+1}{n^2(n-1)} \sum \sum t_{ij} t_{jh} - \frac{1}{n^2} \left[ \sum \sum t_{ij} t_{hj} + \sum \sum t_{ij} t_{ih} \right]$$

Cette formule nous permet d'éliminer les conditions sur les indices  $i, j, h$  et  $l$ , mais reste très coûteuse en temps de calcul.

Une seconde alternative est alors considérée. Elle consiste à prendre une sous-population formée de solutions générées par une heuristique; un échantillon de taille  $n$  est ensuite tiré de cette population. Cette alternative permet d'éliminer les mauvaises solutions.

b/Démarche pour la détermination des bornes par la méthode des moments:

L'approche peut se résumer ainsi:

On a un graphe à  $(n+1)$  sommets (représentation graphique du TSP), le nombre total de circuits hamiltoniens possibles est  $n!$ , on connaît la moyenne et la variance de la population constituée par le nombre total de circuits. Ensuite, nous tirons un échantillon de taille  $m$  suffisamment grande, sa moyenne est distribuée selon la loi normale de moyenne  $T$  et d'écart-type  $s/\sqrt{m}$  (théorème central limite) dans les cas pratiques  $m \geq 30$ .

Pour pouvoir vérifier le bien fondé de cette approche nous avons refait les travaux de GOLDEN et al [5, 6, 7].

Pour mettre en oeuvre la procédure de construction de l'intervalle de confiance décrite ci-dessus .Nous avons eu recours au procédé suivant :

-Constitution de l'échantillon de solutions en utilisant la procédure de génération aléatoire de circuits, puis en appliquant la procédure d'affinage 2-opt aux circuits obtenus.

-Utilisation du logiciel UNIFIT [Annexe 4], version 2.11 /PC Oct.1985 pour l'estimation des paramètres de la loi de Weibull.

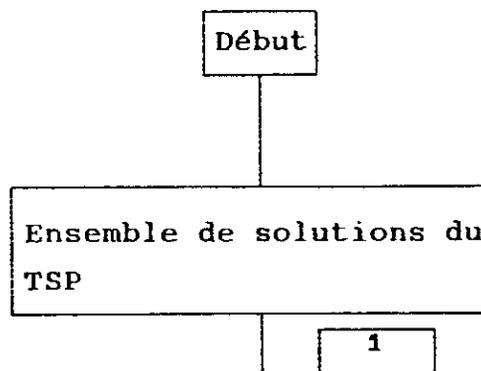
#### IV/DESCRIPTION ET MISE EN OEUVRE DES PROCEDURES.

Nous donnerons dans cette partie les différentes procédures et la façon dont elles sont utilisées.

##### IV.1. PROCEDURE GENERALE:

Elle peut se résumer de la manière suivante:

1°)cas:la population est constituée de tous les circuits hamiltoniens c'est-à-dire elle ne subit aucune modification.



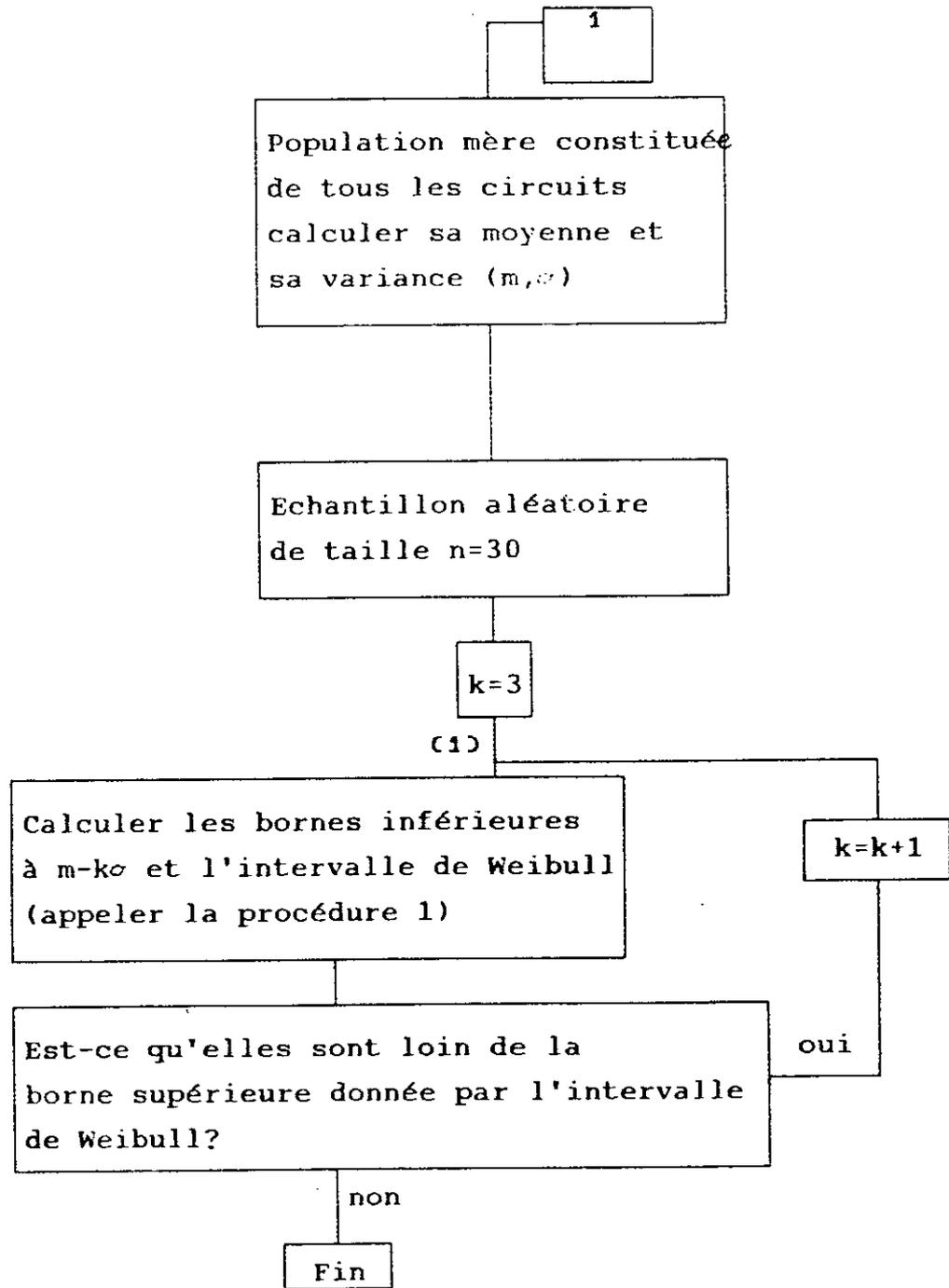


Fig.2: Procédure générale cas 1.

2°)cas:la population est modifiée de manière à n'être constituée que de solutions données par la procédure d'affinage 2-opt population "réduite".

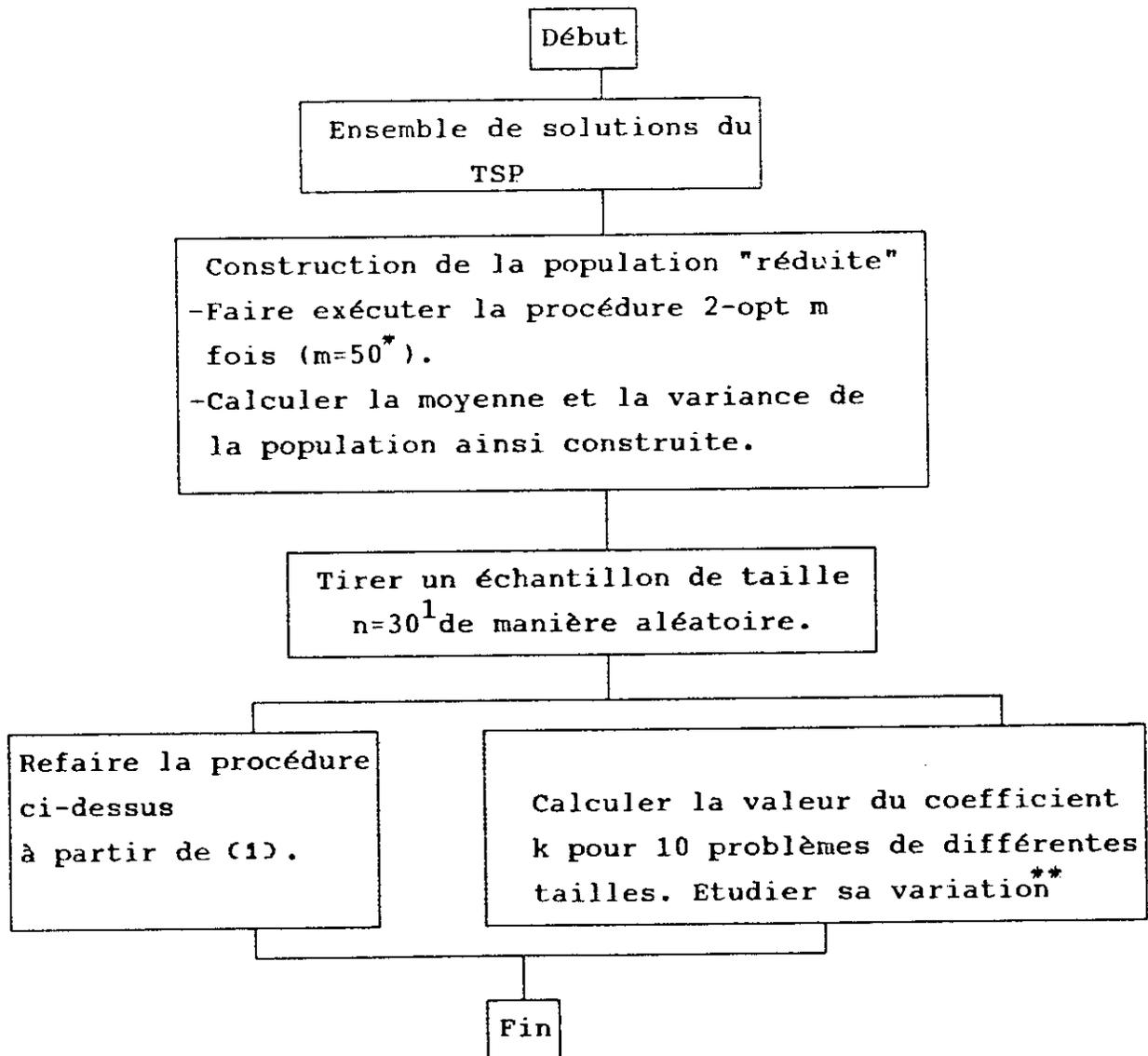


Fig.3:Procédure générale cas 2.

\*:Vu les capacités matérielles disponibles nous nous sommes

contenté, dans un premier temps, de populations de taille 50 avec 10 problèmes seulement.

1/:Le logiciel UNIFIT donne de "bons" ajustements quand la taille de l'échantillon est égale à 30.

\*\* :  $k = (\text{min} - m) / \alpha$ .

Les résultats de cette procédure générale qui fait appel à d'autres procédures seront donnés après avoir passé en revue ces dernières [IV.4].

#### IV.2. CHOIX ET DESCRIPTION DE L'HEURISTIQUE:

##### 1. Description de la procédure r-opt:

Cette procédure rentre dans le cadre des heuristiques de recherche locale "Local Search Heuristic", qui sont les plus efficaces pour l'obtention d'une solution pour le TSP [13,22]. Elle se résume ainsi:

- 1-Commencer avec une tournée initiale (circuit hamiltonien) H.
  - 2-Enlever r arcs de H ,ce qui donne des chemins discontinus (certains peuvent être des sommets isolés ).
  - 3-Reconnecter les chemins de manière à obtenir un autre circuit hamiltonien H'.
  - 4-Calculer  $w(H')$ , la longueur totale de H'.
- Si  $w(H') < w(H)$  ,remplacer H par H' , et répéter le processus.  
Sinon, prendre un autre ensemble de r arcs de H dans le processus d'échange.
- 5-Répéter les opérations jusqu'à ce qu'il n'y ait plus d'amélioration.

Cette procédure d'affinage (r-opt) peut être utilisée pour tout TSP qu'il soit symétrique ou asymétrique.

Les procédures 2-opt et 3-opt sont obtenues en échangeant respectivement 2 arcs ou 3 arcs .plus le nombre d'arcs à échanger

augmente, plus le temps de traitement de la procédure augmente; il s'agit alors de faire un compromis entre l'exactitude de la solution et le temps de calcul.

## 2. Choix de l'heuristique:

Des problèmes ont été générés aléatoirement puis résolus par:

-Construction de tour géant .

-Raffinage de la solution initiale en utilisant 2-opt et 3-opt .

Les résultats obtenus par 2-opt sont meilleurs que ceux donnés par 3-opt [annexe 5]. Ceci confirme ce qui est dit dans l'ouvrage [13]. En raison de cela, nous avons utilisé 2-opt pour générer des solutions.

### Remarque:

Le choix de la tournée initiale est très important, car il influence considérablement la solution finale obtenue en appliquant 2-opt. Toutefois, cette influence n'est pas toujours "perceptible", dans la mesure où, on peut construire des cas où de "bonnes" solutions initiales ne donnent pas de "bonnes" solutions finales. Mais, il est établi, dans le cas général, par ADRABINSKI & SYSLO (1980) et GOLDEN et al. (1980), qu'une "bonne" solution initiale conduit souvent à une "bonne" solution finale [22].

Nous avons choisi de générer la solution initiale de façon aléatoire, au lieu de le faire avec une procédure de construction de tournée (*Farthest insertion* par exemple). Ce choix a été motivé par le travail déjà effectué dans le domaine de l'analyse statistique.

### -Procédure 2-opt:

Soient:

$H = \{x_1, x_2, \dots, x_n\}$  : Le circuit hamiltonien initial.

$X = \{(x_i, x_j)\}$  : ensemble d'arcs de  $H$  substitué par

$Y = \{(y_p, y_q)\}$  s'il y a amélioration de la solution initiale.

$H' = (H - X) \cup Y$  : la nouvelle tournée raffinée.

Il y a  $\frac{n(n-1)}{2} - n = \frac{n(n-3)}{2}$  tournées  $H'$  pour un  $H$  donné.

Pour chacune de celles-ci.

Soit  $\delta$  : l'amélioration réalisée.

$\delta = w(H) - w(H')$        $w(H)$  : longueur du circuit  $H$ .

On cherche parmi les circuits  $H'$ , celui qui maximise  $\delta$ .

Si  $\delta_{\max} \leq 0$ , nous avons la solution finale donnée par 2-opt.

Si  $\delta_{\max} > 0$ , nous utilisons la solution correspondante comme solution initiale, et nous répétons la procédure 2-opt jusqu'à ce que  $\delta_{\max}$  soit non positif.

#### L'Algorithme général :

Begin

  Soit  $H = \{x_1, x_2, \dots, x_n\}$  solution initiale.

  repeat

$\delta_{\max} \leftarrow 0$ ;

    For  $i \leftarrow 1$  to  $(n-2)$  do

      For  $j \leftarrow (i+2)$

    ) to  $n$  or  $(n-1)$  do

      {Le dernier cas est seulement pour  $i=1$ }

      If  $(w(x) + w(x) - w(y) - w(y)) > \delta_{\max}$  then

        Begin  $\delta_{\max} \leftarrow (w(x_i) + w(x_j) - w(y_q) - w(y_p))$ ;

        sauvegarder  $i$  et  $j$  End;

    If  $\delta_{\max} > 0$  then

$H \leftarrow H - \{x_i, x_j\} \cup \{y_p, y_q\}$  ;

  Until  $\delta_{\max} = 0$ ;

End.

#### IV. 3. CHOIX DE L'ECHANTILLON:

Le choix de l'échantillon de solutions a une grande importance car les résultats en dépendent fortement (des solutions loin de l'optimum peuvent donner des intervalles qui ne contiennent pas l'optimum dans la majorité des cas et plus les solutions sont proches de l'optimum plus l'intervalle obtenu a une grande chance de contenir la solution optimale).

Le choix de l'échantillon peut se faire de deux façons:

1°/Echantillon tiré aléatoirement.

2°/Echantillon est construit à partir d'un ensemble de solutions qui font partie du pire des cas, d'autres qui font partie du meilleur des cas et d'autres parmi les plus probables.

Cette dernière présente, cependant, des difficultés qui sont dues au fait qu'il n'y a pas de règles objectives applicables qui permettent de bien choisir le nombre de solutions de chaque catégorie. C'est pourquoi un tirage aléatoire semble plus approprié.

#### IV. 4. IMPLEMENTATION:

La procédure générale, codé en TURBO PASCAL version 6.0, a été implémentée sur un micro-ordinateur de type IBM P.S/2, à l'Ecole Polytechnique d'Alger. Toutes ces procédures ont porté sur des problèmes générés de manière aléatoire.

Une fois que les échantillons sont obtenus, ils sont traités par le logiciel UNIFIT pour deux raisons:

-tester l'hypothèse: le minimum de l'échantillon est distribué selon la loi de Weibull,

-obtenir une estimation, par la méthode du maximum de vraisemblance, des paramètres de la loi.

Les intervalles de confiance pour le minimum sont alors

obtenus, ainsi que d'autres résultats relatifs à la procédure des moments [IV.4].

VI.5. RESULTATS ET INTERPRETATIONS:1) RESULTATS:

Avant de donner les différents résultats, nous allons passer en revue toutes les procédures que nous avons utilisées.

a) Méthode utilisant la loi de distribution de Weibull:

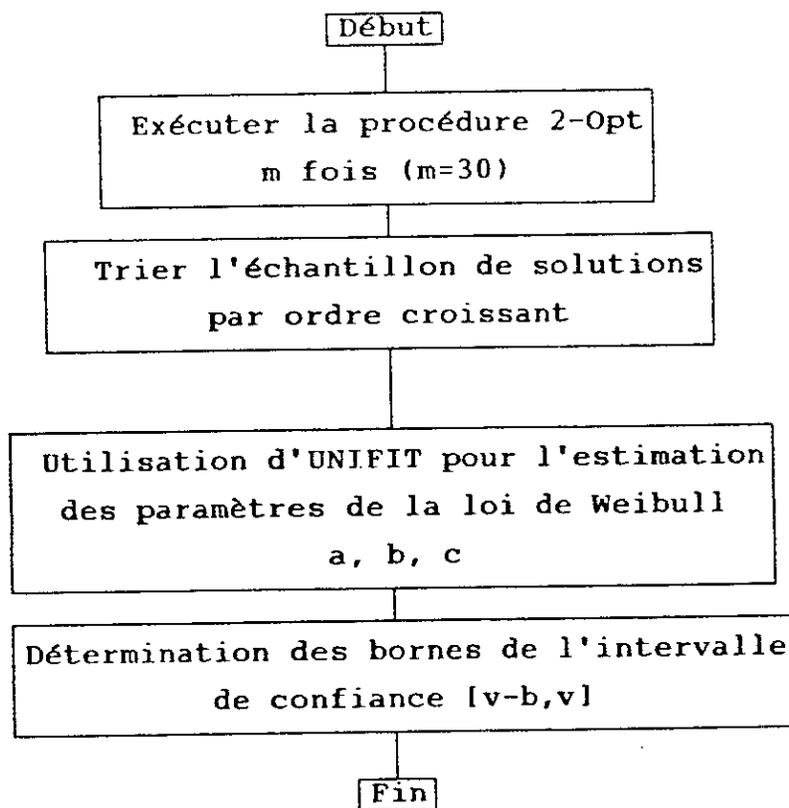


Fig.4: Procédure de construction de l'intervalle de confiance de Weibull.

b) Méthode de détermination des bornes inférieures à  $3\sigma$  et  $4\sigma$  en utilisant les moments de la population:

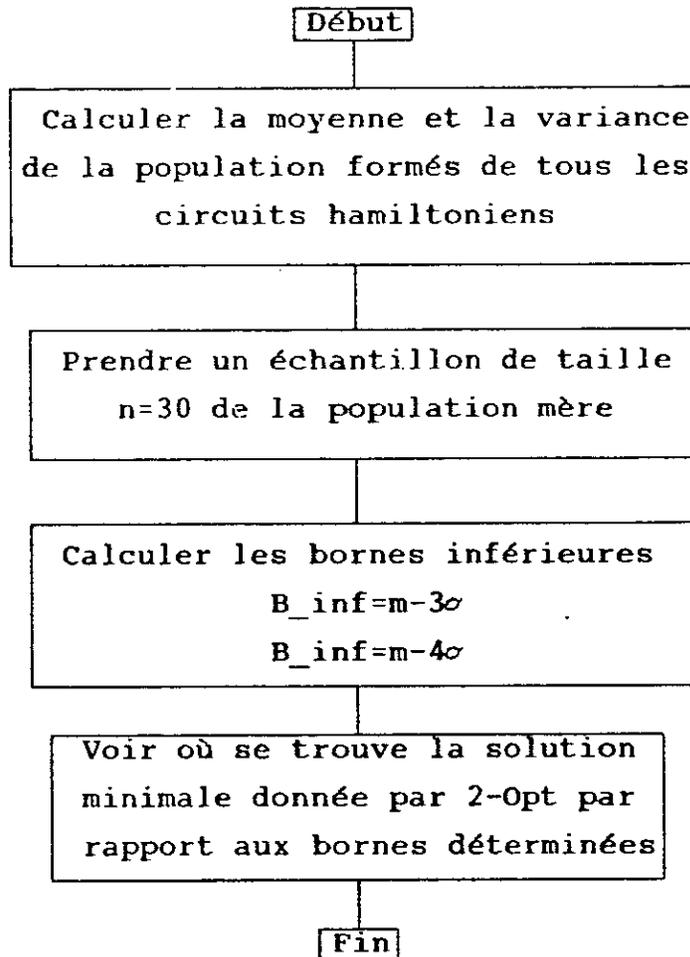


Fig.5: Détermination des bornes inférieures.

c) Méthode de comparaison entre la procédure 1) et 2):

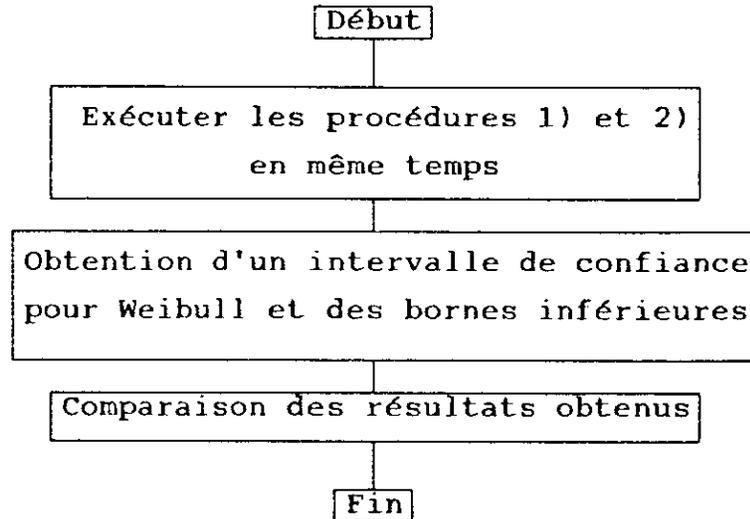


Fig.6: Comparaison entre les procédures 1) et 2)

Une autre approche pour déterminer les bornes inférieures moyennant l'utilisation des moments consiste à changer les individus de la population mère.

Ceci peut être réalisé de la manière suivante:

Nous savons que la population mère est constituée dans sa grande majorité de "mauvais" circuits Hamiltoniens. Cette situation est l'un des paramètres qui provoque un éloignement des bornes supérieures données par l'approche de Golden et al. Par conséquent, si nous arrivons à prendre une autre population où la majorité de ces individus sont de "bons" circuits Hamiltoniens, nous pourrions peut être réduire cet éloignement.

La méthode utilisée consiste à prendre une population composée uniquement de solutions obtenues par la procédure 2-Opt.

d) Méthode utilisant la population "réduite":

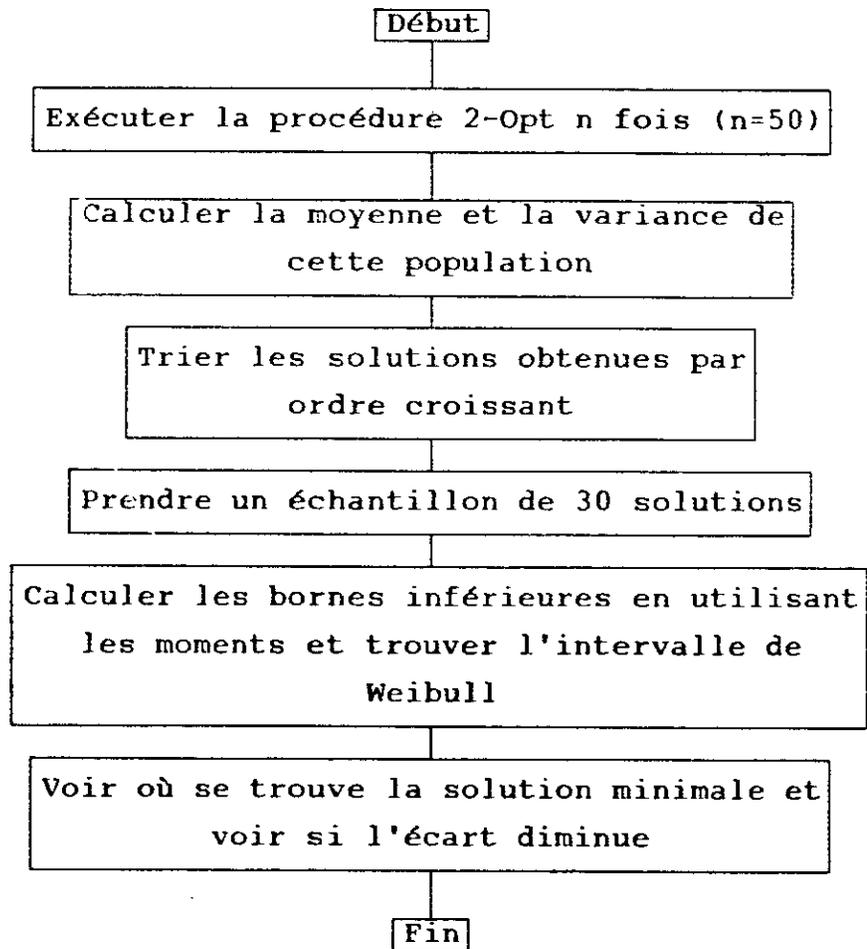


Fig.7:procédure utilisant la population "réduite".

les résultats sont donnés pour toutes les procédures ci-dessus.

Procédure 1: Détermination de l'intervalle de confiance en utilisant la loi de Weibull

Tableau: Résultats de calcul, niveau de confiance  $(1-\exp(-S))100\%$

$S=90$  (\*)

Taille	Minimum donné par 2-opt	Estimation des paramètres de la loi de Weibull			Borne inférieure v-b	Borne sup v
		a	b	c		
20	31937.70	31905.30	1099.56	1.13	[30838.14,	31937.70]
30	37283.80	37283.30	1737.80	0.95	[35543.00,	37283.80]
40	42840.20	42833.80	2618.96	1.19	[40221.24,	42840.20]
50	54906.30	54828.00	3062.65	1.26	[51843.64,	54906.30]
60	76194.33	76193.50	4750.25	1.32	[71444.08,	76194.33]
70	83943.36	83917.80	4896.00	1.70	[79047.36,	83943.36]
80	83864.82	80753.00	6514.40	1.19	[72350.42,	83864.82]
90	91297.36	91236.70	4175.40	1.48	[87121.86,	91297.36]
100	99239.18	99218.60	4271.13	1.45	[94968.05,	99239.18]

(\*) Pour  $S=30$ ,  $1-\exp(-S)$  est pratiquement égal à 1.

Fig.8: Résultats de la procédure 1.

Procédure 2: Détermination des bornes inférieures en utilisant la

moyenne et la variance de la population:

Tableau2:Bornes inférieures lorsque la taille des échantillons est égale à 50

Taille	Longueur moyenne	Ecart-type	Borne inférieure moy- $3\sigma/\sqrt{50}$	Borne inférieure moy- $4\sigma/\sqrt{50}$
30	72107.50	7913.50	68750.09	67630.95
40	102109.70	9977.73	97875.87	96464.81
50	116751.23	9533.64	112706.45	111358.19

Fig.9:Résultats de la procédure 2.

Procédure3: Comparaison entre les résultats de la procédure 1 et ceux de la procédure 2 .

Tableau 3: Résultats comparatifs de la procédure 1 et 2

Taille du problème		30	40	50
Moyenne de la population (m)		72107.50	102109.07	116751.23
Ecart-type de la population ( $\sigma$ )		07913.50	009977.73	009533.64
Minimum donné par 2-opt (v)		14499.80	015622.20	016286.90
Bornes inférieures	$b1=m-3\sigma$	48367.00	072175.87	088150.31
	$b2=m-4\sigma$	40453.50	062198.15	078616.67
	$b3=m-3/\sigma\sqrt{50}$	68750.09	097875.87	112706.45
	$b4=m-4/\sigma\sqrt{50}$	67630.95	096464.81	111358.19
Ecart entre la borne inférieure et v	v-b1	33867.20	56553.68	71863.41
	v-b2	25953.70	46575.95	62329.77
	v-b3	54250.28	82253.67	96419.55
	v-b4	53131.15	80842.61	95071.29
$k1=(m-v)/\sigma$		07.28	08.66	10.53
$k2=(m-v)/\sigma\sqrt{50}$		51.47	61.29	74.51

Fig.10: Résultats de la procédure 3.

Procédure4: Détermination des bornes inférieures en utilisant les

moments de la population "réduite" .

Tableaux: Résultats expérimentaux .

N	k*	Intervalle de confiance de Weibull		Bornes inférieures	
		v-b	v	Borne inf. 3 $\sigma$	Borne inf. 4 $\sigma$
30	2.92	[14396.60,	14638.00]	[14059.11	[13766.25
40	2.48	[16029.41,	16653.60]	[16055.28	[15663.04
50	2.30	[17438.72,	18509.20]	[18249.20	[17875.84
60	1.58	[19135.68,	20044.30]	[19333.47	[18833.43
70	1.89	[23032.38,	24096.20]	[25019.60	[24528.17
80	2.97	[21731.07,	23821.70]	[23804.51	[23227.91
90	2.22	[24031.60,	25537.00]	[25064.90	[24454.81
100	1.53	[26421.07,	27536.60]	[28554.00	[27885.77

\*:k est la valeur de l'expression  $(n-k\sigma)$ , dans ce cas, il représente la valeur à prendre pour avoir la solution minimale donnée par 2-Opt.

Fig.11: résultats de la procédure 4.

2) INTERPRETATIONS:

-----  
Procédure 1: construction de l'intervalle de confiance en utilisant la loi de Weibull.  
-----

En analysant les résultats du tableau 1, nous pouvons dire ce qui suit:

La solution optimale du TSP, qui est dans le pire des cas égale au minimum donné par 2-Opt, se trouvera nécessairement dans deux cas possibles:

1-Cette solution optimale appartient à l'intervalle de confiance

2-Elle est inférieure à  $v-b$  (borne inférieure de l'intervalle de confiance).

Mais comme la probabilité que la solution optimale appartienne à l'intervalle de confiance est égale à  $1-\exp(-S)$  ( $S=30$ ), on peut être pratiquement sûr que c'est le cas 1 qui se présentera dans la majorité des problèmes que nous aurons à traiter.

Le rapport  $b/v$ , qui mesure la performance de l'heuristique utilisée (2-Opt dans notre cas) représente la largeur relative de l'intervalle de confiance et varie entre 4% et 6%.

Il représente aussi "l'erreur relative" commise en prenant la borne supérieure ( $v$ ) comme la longueur minimale du TSP, à défaut de la solution optimale.

Nous avons aussi vérifié l'hypothèse nulle suivante, moyennant l'utilisation du logiciel UNIFIT:

le minimum d'un échantillon de taille  $m$  tiré d'une population, quand  $m$  augmente, est distribué selon une loi de Weibull.

Le risque de se tromper ayant été fixé à 0.05 pour tous les problèmes que nous avons traité. Toutes les hypothèses ont été acceptées.

En pratique, l'utilisateur de cette démarche se trouvera dans deux situations:

a-Pour résoudre son TSP, il pourra prendre la borne supérieure comme étant sa solution ou simplement la solution donnée par 2-Opt (sans construction de l'intervalle dans ce dernier cas).

b-Prendre une valeur appartenant à l'intervalle de confiance.

L'utilisation du cas b- lui donnera une meilleure solution que le cas où il prendra la solution donnée par 2-Opt.

Cette alternative est intéressante mais demandera un temps de calcul sensiblement élevé. Il s'agira alors de faire un compromis entre l'exactitude de la solution et le temps calcul nécessaire pour son obtention. Tout dépendra alors de l'objectif de l'utilisateur.

-----  
Procédure 2: détermination des bornes inférieures en utilisant les moments de la population.  
-----

Tenant compte du temps de calcul requis pour la procédure 2, nous n'avons généré que trois problèmes de tailles 30, 40 et 50. Les résultats obtenus pour  $k=3$  et 4 sont assez loins du minimum de l'échantillon formé par les solutions données par 2-Opt.

Le coefficient  $k=(\min-m)/\sigma$  nous donne la valeur qu'il faudrait prendre, et pour laquelle la borne inférieure coïncide avec le minimum de l'échantillon. Ce coefficient augmente de

façon aléatoire en fonction de la taille du problème.

-----  
Procédure 3: comparaison entre l'intervalle de confiance de Weibull et les bornes inférieures de la méthode des moments  
-----

Les bornes supérieures de l'intervalle de Weibull sont inférieures aux bornes données par la méthode des moments.

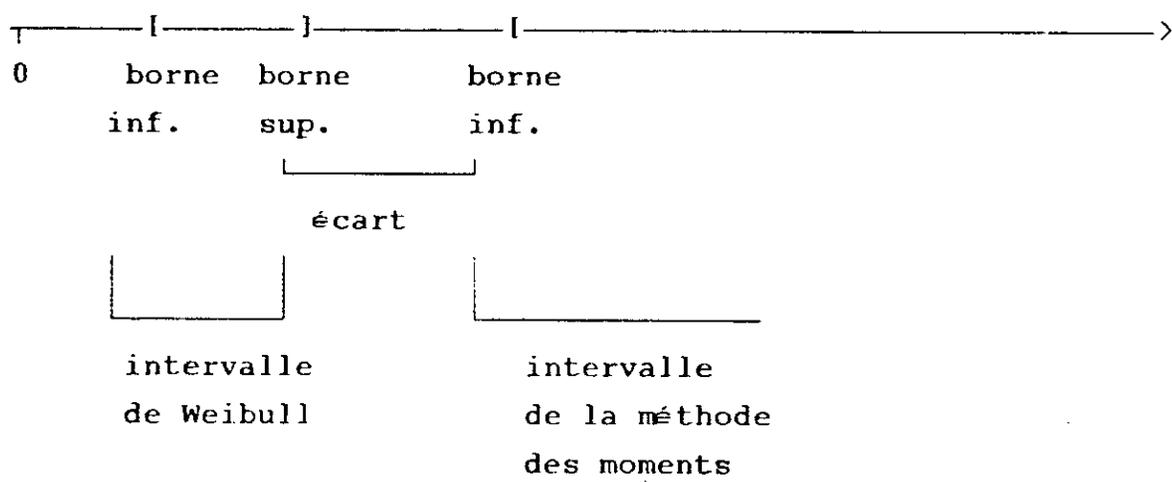


Fig 12:représentation schématique de la comparaison.

L'écart entre ces deux bornes augmente en fonction de la taille du problème pour une valeur du coefficient k fixé à l'avance.

-----  
 Procédure 4:détermination des bornes inférieures en utilisant les moments de la population "réduite"  
 -----

Le coefficient k étant fixé (3 et 4), cette procédure améliore les bornes inférieures (réduit l'écart cité ci-dessus).

CONCLUSION

L'objectif assigné à ce modeste travail n'était pas de trouver une solution optimale au problème du voyageur de commerce (TSP), mais plutôt trouver, par une méthode statistique, l'intervalle de confiance où cette solution pourrait se trouver.

Pour le faire, nous avons pensé à la nature de la population formée par les circuits Hamiltoniens (ensemble de solutions possibles du TSP). Une idée classique de la théorie statistique nous vient alors à l'esprit: c'était le calcul de la moyenne et de la variance de cette population.

Une fois ce travail réalisé, nous avons construit l'intervalle ci-dessus cité. Nous avons, par la suite, essayé de valider cette approche utilisant les moments. Cette étape a été réalisée en s'aidant d'une autre approche dans le domaine de l'optimisation statistique. Cette dernière a été utilisée par plusieurs auteurs parmi lesquels GOLDEN apparaît en relief.

Golden et al. ont généré leurs solutions initiales, améliorées ensuite par la procédure 2-Opt, en utilisant le logiciel RANPER. Ce dernier n'est pas disponible à notre niveau. Chose qui nous a obligé à refaire tout le travail pour pouvoir comparer les résultats obtenus par ces deux approches (la notre et celle de Golden et al.).

Notre méthode consiste à considérer deux populations, l'une composée de toutes les solutions du TSP et l'autre composée uniquement de solutions obtenues en déroulant la procédure 2-Opt 50 fois. A partir de ces deux populations, nous tirons des échantillons de manière aléatoire.

Nous avons appliqué les différentes procédures précédemment décrites sur les deux types de populations en fixant la valeur du coefficient  $k$  à 3 et 4 écart-types pour l'approche des moments. Les résultats obtenus pour le deuxième type de population sont meilleurs comparés à ceux obtenus pour le premier type de population.

Nous avons mis en oeuvre une procédure de détermination du coefficient  $k$ . Ceci a été réalisé pour deux cas de figures:

-pour une même taille de problème, nous avons généré aléatoirement 10 problèmes différents.

-pour une même taille, nous avons considéré uniquement 1 problème généré aléatoirement.

Pour ces deux cas, la valeur du coefficient  $k$  fluctue de façon aléatoire [annexe 4].

Il est, par conséquent difficile, de se prononcer sur la valeur que ce coefficient prendra pour avoir le minimum.

Ces méthodes d'estimation sont limitées en pratique car elles sont gourmandes en temps de calcul comparées aux heuristiques déjà existantes. Néanmoins, elles restent un outil pour l'évaluation de la performance des heuristiques et des solutions qui en ressortent.

Il serait souhaitable de continuer l'étude pour essayer de déterminer la loi de distribution, non pas de l'échantillon, mais de la population mère.

Une autre alternative serait d'étudier la variation du coefficient  $k$ , mais ceci supposerait la disponibilité de moyens de prospection (informatiques surtout) suffisamment puissants.

Ce travail représente une contribution théorique dans le domaine de l'optimisation combinatoire, domaine qui ne cesse de connaître un développement...combinatoire!

Les recherches sur le TSP sont loin d'être closes. La résolution optimale de ce problème supposera un développement, plus que jamais, nécessaire de l'algorithmique et de l'informatique. Mais sa résolution ouvrira d'intéressantes perspectives pour tous les problèmes dont il constitue une composante principale. Et il cessera peut être alors de hanter un bon nombre de chercheurs.

Nous espérons, enfin, que cette démarche sera appliquée à d'autres problèmes d'optimisation combinatoire.

BIBLIOGRAPHIE:

- [1] J.L.AARTHUR & J.O.FRENDEWEY  
"Generating TSPs with known optimal tours"  
Operational Research Society.Vol 39,n°2,pp 153-159,(1988).
- [2] R.BELLMAN ,K.L.COOK & J.A.LOCKETT  
"Algorithms,graphics & computers".  
Vol 62 ,Academic press,(1970).
- [3] R.BENAYOUM, P.T.NGHIEM & B.ROY  
"Un mod le d'implantation et de cheminement".  
Doc.SEMA,(mai 1966).
- [4] J.BERGE  
"Graphes et leurs applications"  
Bordas,(1958).
- [5] B.GNEDENKO  
"Sur la distribution limite du terme maximum d'une série  
aléatoire"  
Ann.Math.44, 3, (1943).
- [6] B.L.GOLDEN  
"A statistical approach to the TSP".  
Network 7,pp 209-225,(1977b):
- [7] B.L.GOLDEN  
"Point estimation of a global optimum for large  
combinatorial problems"  
Comm.statist. B-simulation comp.7, pp 361-367, (1978).

- [8] B.L.GOLDEN & F.B.ALT  
*"Interval estimation of a global optimum for large combinatorial problems"*  
Naval Res. Logist. Quart.26, pp 69-77, (1979).
- [9] M.GONDRAN & M.MINOUX  
*"Graphes et algorithmes"*  
Ellipse, (1979).
- [10] E.GUMBEL  
*"Statistics of extremes"*  
Columbia University press, N.Y. (1958).
- [11] D.E.KNUTH  
*"Sorting & searching : The art of computer programming"*  
Vol 1,2 et 3, Addison Wesley, (1973).
- [12] A.LANGEVIN, F.SOUMIS & J.DESROSIERS  
*"Classification of TSP formulations"*  
publication de l'université de montréal, GERAD, HEC, (1988).
- [13] E.L.LAWLER, J.K.LENSTRA, A.H.G.RINNOY KAN & D.B.SHMOYS  
*"The traveling salesman problem: A guided tour of combinatorial optimisation"*  
John Wiley & sons Ltd, (1987).
- [14] J.D.C.LITTLE  
*"An algorithm for the traveling salesman problem"*  
O.R.S.A.11, (1963).

- [15] M.LOS & C.LARDINOIS  
*"Combinatorial programming, statistical optimization and the optimal transportation network problem"*  
Transportation Research, part B.16, pp 89-124, (1982).
- [16] M.MINOUX,G.PARTNIK  
*"Graphs, algorithms & logiciels"*  
DUNOD Info., BORDAS, (1986).
- [17] A.M.MOOD  
*"Introduction to the theory of statistics"*  
Mc graw hill, (1975).
- [18] A.PAPOULIS  
*"Probability, random variables & stochastic process"*  
Mc Graw Hill, (1985).
- [19] D.J.ROSENKRANTZ ,R.E.STERANS & P.M LEWIS II  
*"An analysis of several heuristics for the TSP "*  
SIAM. J. comput. Vol 6, n°3, (1977).
- [20] M.SAKAROVITCH  
*"Optimisation combinatoire, programmation discrète".*  
Hermann, (1984).
- [21] M SARI  
*"Investigation des modèles à coûts unitaires variables dans le cas d'un système de distribution à dépôt unique ou à plusieurs dépôts "*  
Thèse de Magister en G.Industriel,Ecole Polytechnique d'Alger, ALGERIE, (1990).

[22] M.M.SYSLO N.DEO & J.S.KOWALIK

*"Discrete optimisation algorithms"*

Printice Hall, (1983).

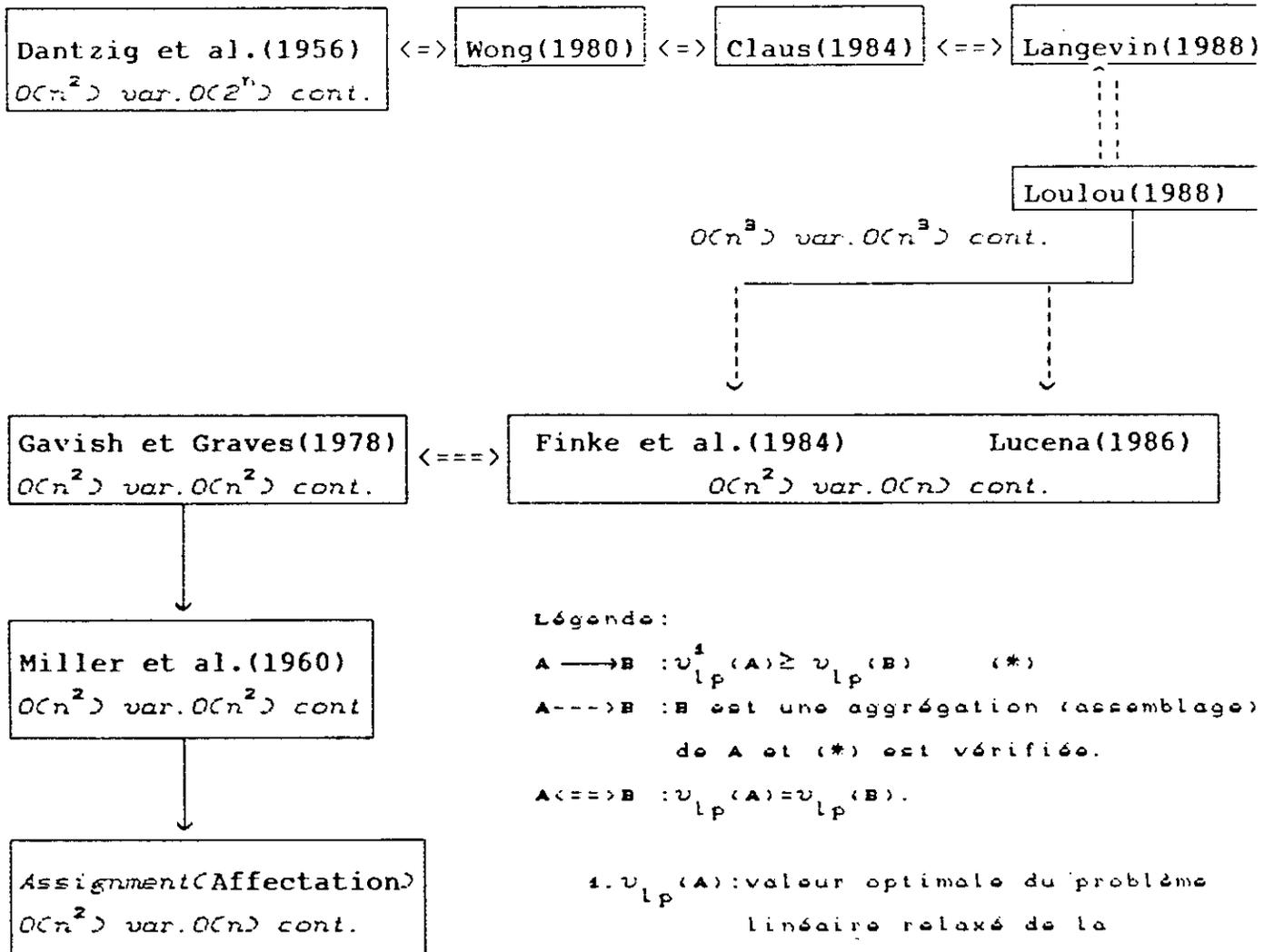
[23] L.VILLENEUVE

*"Introduction to UNIFIT "*

G.Industriel, Ecole Polytechnique de Montréal, CANADA, (1988).

## Annexe 1:

## Classification des relaxations du TSP.



## Annexe 2:

Résultats donnés par GOLDEN & ALT(1979) pour un intervalle de confiance  
à 100 [1-exp(-s)]%\*

Pb.N°	v	$\tilde{a}$	$\tilde{b}$	$\tilde{c}$	$[v-\tilde{b},v]$	valeur de la sol.opt.(X0)	X0 [v- $\tilde{b}$ ,
24	21518.99	21454.77	926.36	1.76	[20592.63,21518.99]	21282	Oui
25	22816.55	22754.95	856.92	1.92	[21959.63,22816.55]	22141	Oui
26	20971.54	20763.87	1260.75	2.31	[19710.79,20971.54]	20749	Oui
27	21807.23	21779.19	837.82	1.51	[20969.41,21807.23]	21294	Oui
28	22382.39	22190.20	1415.64	2.62	[20966.75,22383.39]	22068	Oui
318	41415.00	41363.38	609.58	1.96	[40805.42,41415.00]	41345	Oui

\*.Pour ces problèmes,s=50 et  $\exp(-50)=192874.10^{-22}$ .Ceci implique que  
1-exp(-s) est pratiquement 1.

## Annexe 3:

Mesure de l'efficacité des algorithmes (deux méthodes):

La théorie de la complexité des algorithmes est née vers la fin des années 60, de la nécessité de disposer d'une mesure comparative objective universelle de l'efficacité des algorithmes. Il existe deux méthodes essentielles:

1) Mésure expérimentale de la complexité:

La première idée qui vient à l'esprit, pour déterminer la complexité d'un algorithme, consiste à recourir à la méthode expérimentale: obtenir les temps de calcul obtenus sur des séries de problèmes tests de taille suffisamment importante pour trouver (par exemple par une méthode des moindres carrés) une fonction  $f$  permettant de rendre compte, au mieux, des observations effectuées. Cependant, une telle méthode ne permet généralement pas la détermination précise de la complexité, car elle souffre de deux défauts majeurs:

- On ne peut jamais être sûr d'avoir atteint des tailles de problèmes suffisamment grandes pour se trouver dans une zone où le comportement de la fonction  $t_n$  (la croissance asymptotique du temps de calcul en fonction de la taille du problème) est suffisamment proche de son comportement asymptotique;

- la complexité des algorithmes est une notion relative au pire des cas, c'est-à-dire est censée rendre compte du comportement de l'algorithme pour les problèmes qui seront les plus difficiles à traiter. En fait, il n'existe que peu de couples (problème, algorithme) pour lesquels on sait effectivement définir pour chaque taille de problème, les cas les plus difficiles à résoudre (c'est-à-dire ceux qui nécessitent le temps de calcul le plus élevé). De plus, la mise en évidence d'un pire cas est généralement un problème beaucoup plus difficile que celui de la

détermination directe de la complexité (par la méthode qui sera exposée au 2<sup>e</sup> paragraphe ci-dessous).

Remarquons que, par opposition à la notion de complexité apparaissant dans le chapitre I paragraphe I.5.1., laquelle se réfère implicitement au pire des cas, la méthode expérimentale, qui consiste à procéder à des observations du temps de calcul sur un échantillonnage de problèmes, conduit plutôt à mesurer ce que l'on pourrait appeler la "complexité en moyenne" ou encore "au sens du cas moyen". (on peut donner une signification rigoureuse à cette notion de "cas moyen" à condition de préciser la notion de tirage au hasard d'un problème particulier dans un ensemble de problèmes, autrement dit, de définir une distribution de probabilité sur l'espace des données d'un problème).

L'utilisation d'une mesure expérimentale de la "complexité en moyenne" peut se justifier, dans un certain nombre de situations, par exemple, la détermination de la complexité théorique est impossible ou, surtout, lorsque le comportement de l'algorithme dans le "cas moyen" est extrêmement différent de celui pour le pire des cas. Dans cette dernière situation, la mesure "en moyenne" représente souvent beaucoup mieux le comportement réel de l'algorithme.

## 2) Détermination de la complexité par décompte du nombre d'opérations élémentaires:

Pour montrer comment la détermination de la complexité peut se faire autrement, et de façon plus rigoureuse, que par la méthode expérimentale, nous commencerons par étudier un exemple:

soit une matrice  $A=(a_{ij})$  d'ordre  $N$  et à coefficients 0 ou 1. Soit  $A'$  la matrice déduite de  $A$  en ajoutant  $N$  termes diagonaux égaux à 1. Élevant la matrice  $A'$  à la puissance  $N-1$  dans l'algèbre obtenue:

- en remplaçant l'addition ordinaire par l'opération \*

définie par  $a \oplus b = \text{Max}\{a, b\}$ ,  $a, b \in \{0, 1\}$  ;

- en remplaçant la multiplication ordinaire par l'opération définie par  $a \otimes b = \text{Min}\{a, b\}$ ,  $a, b \in \{0, 1\}$  ;

On sait que le calcul de  $A^{N-1}$  peut se faire en au plus  $2\lceil \log_2(N-1) \rceil$  multiplications de deux matrices  $(N, N)$  (ce résultat est aisément obtenu en décomposant  $N-1$  suivant la base 2). Pour effectuer le produit de deux matrices quelconques  $A$  et  $B$ , on doit calculer :

$$C_{ij} = \sum_{k=1, \dots, N}^{\oplus} a_{ik} \otimes b_{kj} \quad \forall i, j \quad (1)$$

où le signe  $\sum$  désigne la sommation au sens de l'opération  $\oplus$  définie plus haut.

La complexité d'un tel algorithme pour le produit de deux matrices  $(N, N)$  est facile à évaluer. En effet, pour chaque terme  $(i, j)$ , la formule (1) nécessite  $N$  opérations  $\oplus$  et opérations  $\otimes$ . Convenons alors d'appeler *opération élémentaire* l'ensemble constitué par une opération  $\otimes$  sur deux opérandes pris dans  $\{0, 1\}$ , suivie par une opération  $\oplus$  sur deux opérandes pris dans  $\{0, 1\}$ . On notera que, quel que soit le calculateur utilisé, le temps mis pour l'exécution de cette opération est une constante qui ne dépend pas de la taille du problème traité. Comme elle ne dépend que de la machine  $m$  utilisée, nous la noterons  $c(m)$ .

L'algorithme de multiplication matricielle nécessite  $N$  opérations élémentaires pour le calcul de chaque terme de la matrice produit, donc en tout  $N^3$  opérations élémentaires. Comme le nombre maximal de produits de matrices à effectuer est, dans le pire des cas,  $2\lceil \log_2(N-1) \rceil$  on en déduit que cette algorithme nécessite (dans le pire des cas)  $2c(m)N^3\lceil \log_2(N-1) \rceil$  unités de temps. D'après la définition de la complexité, cette dernière est alors obtenue en "oubliant" le terme constant  $2c(m)$  qui ne dépend que de la machine considérée et l'on déduit :

La complexité de cette algorithme est de l'ordre

$O(N^3 \log_2 (N-1))$ .

En toute rigueur, pour une cohérence parfaite avec la définition sur la complexité, il faudrait exprimer cette dernière en fonction de la taille  $n$  du problème qui est dans ce cas  $n=N^2$ . La complexité de cette algorithmme est  $O(n^{3/2} \log_2 n)$ .

L'exemple ci-dessus constitue une illustration simple, mais caractéristique, de la façon dont la complexité peut être obtenue en procédant à une analyse fine de l'algorithme et au décompte détaillé du nombre d'opérations nécessaires pour son fonctionnement, en oubliant jamais de se placer systématiquement dans le pire des cas qui puisse se présenter. Il est de pratique courante de considérer chacune des opérations arithmétiques comme des opérations élémentaires.

Il existe néanmoins des situations où elle ne se justifie que parce qu'on admet implicitement que l'on se restreint à une sous-classe de problèmes: ceux pour lesquels les valeurs numériques que l'on manipule sont représentables à l'aide d'un nombre maximal fixe d'éléments binaires ; ou encore, ceux pour lesquels la perte de précision résultant du fait de restreindre la représentation des valeurs numériques à un nombre fixe donné d'éléments binaires n'affecte pas significativement le résultat, du point de vue de l'utilisation qui en est faite.

Annexe 4:  
Description d'UNIFIT

-UNIFIT est un logiciel en mode interactif qui permet d'adapter un échantillon de données qu'on a collecté expérimentalement, à une distribution de probabilité théorique.

-Son avantage est qu'il permet de trouver la meilleure représentation qui s'adapte à un cas, et ceci en peu de temps. Chose qui est souvent négligée à cause du temps requis.

-UNIFIT est capable d'adapter un échantillon à l'une des 13 distributions continues ou des 5 distributions discrètes suivantes:

\*Lois continues non négatives: Exponentielle, Gamma, Inverse gaussian, Lognormale, Weibull, Pearson type 5, Pearson type 6.

\*Lois continues (univers non borné): Extrem value type A ; extrem value type B, Logistic, Normale.

\*Lois continues (univers borné): Uniform, Béta.

\*Lois discrètes: Binomiale, Géométrique, Binomiale négative, Poisson, Uniforme.

-Les fonctions de UNIFIT sont regroupées en sept rubriques.

1°/Operational characteristics.

2°/Sample maintenance.

3°/Description sample summaries.

-Histogram.

-Quantil summary and box plot.

-Indépendance plot.

4°/Model specification.

-Non negative continuous models

-Unbounded continuous models.

-Bounded continuous models

-List specified models.

-Parameters confidence intervals.

5°/Heuristics for the model choice.

-Frequency comparison.

-Density/histogram overplot.

-Cumulative frequency comparison.

-Q-Q plot.

-P-P plot.

-Relative discrepancies of probability plots.

-Model moment comparison.

6°/Testing hypothesis.

-Chi square goodness of fit test.

-KOLMOGOROV-SMIRNOV goodness of fit test

-ANDERSON-DARLING goodness of fit test.

-Model test comparison.

7°/Characteristics of fitted distributions.



## Annexe 5:

N°	Longueur T.aléatoire	Longueur 2-Opt	Longueur 3-Opt	3-Opt-2-Opt
1, 20,	3845.38,	1269.71,	2755.82,	1486.11
1, 20,	3607.79,	1204.59,	2030.38,	825.80
1, 20,	3556.20,	1204.59,	1881.36,	676.77
1, 20,	3849.89,	1204.59,	2539.05,	1334.46
1, 20,	3113.96,	1197.31,	2014.63,	817.32
2, 20,	3526.77,	1314.06,	2252.56,	938.50
2, 20,	2989.82,	1242.73,	2070.76,	828.03
2, 20,	3168.40,	1271.53,	1883.69,	612.16
2, 20,	4385.81,	1341.84,	1979.11,	637.27
2, 20,	3335.34,	1289.50,	1533.26,	243.76
3, 20,	3599.28,	1190.15,	1698.17,	508.02
3, 20,	3465.06,	1272.65,	2225.30,	952.65
3, 20,	3934.14,	1272.65,	2076.00,	803.34
3, 20,	3833.23,	1198.43,	2098.07,	899.64
3, 20,	3536.00,	1255.96,	1918.29,	662.33
4, 20,	2913.89,	1180.27,	1803.38,	623.11
4, 20,	3493.07,	1180.27,	1827.77,	647.50
4, 20,	3611.32,	1244.77,	2445.24,	1200.48
4, 20,	3159.50,	1253.35,	1978.23,	724.88
4, 20,	4026.77,	1271.23,	1846.17,	574.94
5, 20,	3327.82,	1307.20,	2335.88,	1028.68
5, 20,	4027.50,	1304.59,	2495.61,	1191.02
5, 20,	3320.39,	1426.22,	2193.04,	766.82
5, 20,	3780.69,	1301.38,	2336.33,	1034.95
5, 20,	3819.74,	1276.77,	2599.54,	1322.77
6, 20,	4055.25,	1245.82,	2775.16,	1529.34
6, 20,	3187.10,	1260.16,	2081.03,	820.88
6, 20,	4046.54,	1211.09,	2618.12,	1407.03
		chacun.		
6, 20,	4037.73,	1211.09,	2500.68,	1289.59
6, 20,	3512.36,	1210.76,	1815.13,	604.38
7, 20,	3269.36,	1141.23,	1855.73,	714.49
7, 20,	3167.63,	1225.37,	1812.26,	586.89
7, 20,	3747.23,	1141.23,	1728.64,	587.40
7, 20,	2852.17,	1234.23,	1864.23,	630.00
7, 20,	3357.05,	1216.87,	1797.65,	580.79
8, 20,	3125.95,	1377.80,	2147.62,	769.81
8, 20,	5265.44,	1285.00,	2873.88,	1588.87
8, 20,	4143.98,	1344.44,	1961.98,	617.54
8, 20,	4935.63,	1285.00,	3429.00,	2143.99
8, 20,	4668.58,	1293.11,	2648.66,	1355.55
9, 20,	3551.12,	1258.66,	2577.92,	1319.26
9, 20,	3267.20,	1195.42,	1859.02,	663.60
9, 20,	3788.00,	1187.46,	2307.31,	1119.85
9, 20,	3266.17,	1226.59,	2201.99,	975.40
9, 20,	3150.55,	1214.95,	2235.16,	1020.22
10, 20,	4804.17,	1322.44,	2063.90,	741.46
10, 20,	4062.22,	1349.93,	1881.45,	531.51
10, 20,	3749.64,	1318.92,	1994.51,	675.59
10, 20,	4262.42,	1386.95,	2516.45,	1129.50
10, 20,	3221.94,	1388.54,	2336.22,	947.68

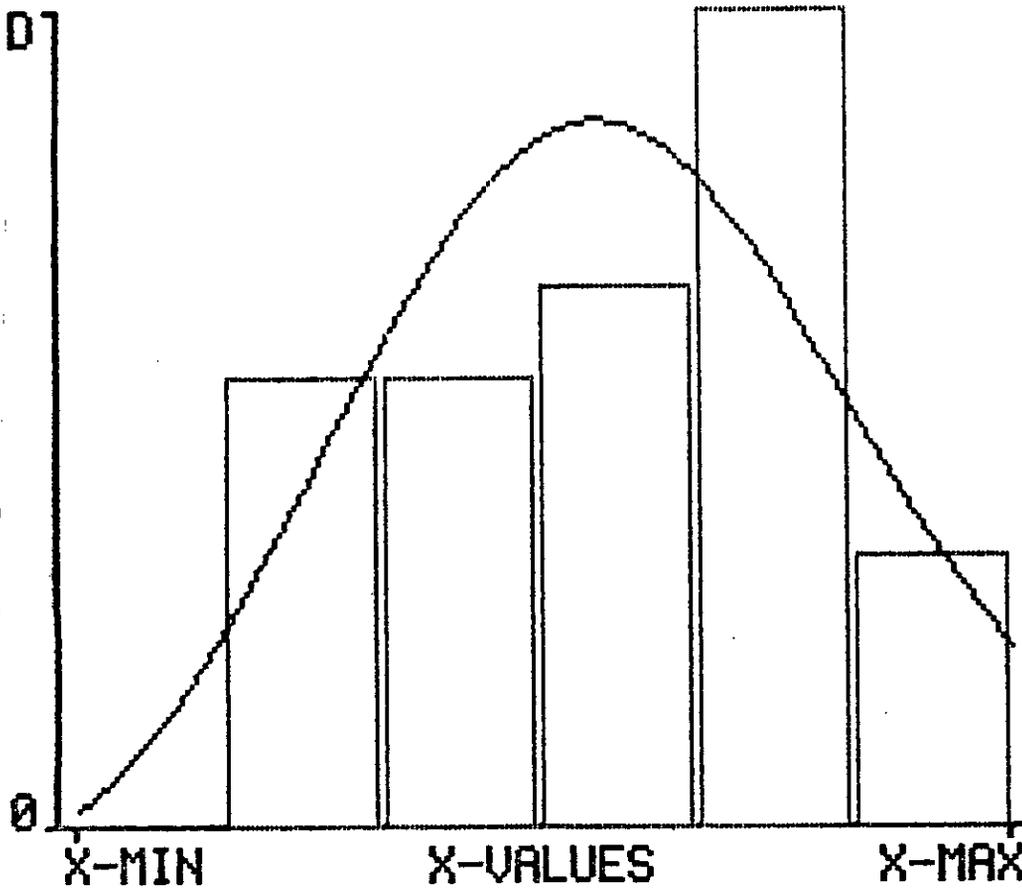
Résultats comparatifs  
entre les procédures  
2-Opt et 3-Opt pour  
5 problèmes de taille 20

## Annexe 5:(suite).

1, 40,	6453.87,	1711.35,	3196.59,	1485.24
1, 40,	7351.91,	1562.26,	3349.55,	1787.29
1, 40,	6607.94,	1466.64,	4158.12,	2691.48
1, 40,	6747.20,	1511.36,	3344.44,	1833.08
1, 40,	6736.12,	1576.76,	3288.12,	1711.36
2, 40,	5356.88,	1629.63,	3370.83,	1741.20
2, 40,	6206.11,	1663.04,	2958.64,	1295.60
2, 40,	5448.69,	1725.16,	3234.15,	1508.99
2, 40,	5661.23,	1594.47,	3406.38,	1811.91
2, 40,	6444.55,	1613.19,	3675.07,	2061.88
3, 40,	7331.92,	1788.50,	3015.41,	1226.91
3, 40,	6870.94,	1606.56,	3482.90,	1876.34
3, 40,	6928.27,	1755.06,	3219.45,	1464.39
3, 40,	6998.43,	1726.52,	4583.94,	2857.43
3, 40,	6587.36,	1687.82,	3182.46,	1494.64
4, 40,	5934.32,	1684.06,	2916.82,	1232.76
4, 40,	5375.46,	1931.95,	2847.46,	915.51
4, 40,	6127.69,	1825.81,	3624.19,	1798.37
4, 40,	6480.16,	1747.56,	2815.57,	1068.01
4, 40,	5660.03,	1657.79,	3673.03,	2015.24
5, 40,	7617.44,	1867.71,	4789.41,	2921.70
5, 40,	7448.72,	1638.89,	3912.85,	2273.96
5, 40,	6594.98,	1661.43,	3772.07,	2110.65
5, 40,	6629.54,	1665.62,	3325.79,	1660.17
5, 40,	6526.97,	1652.08,	3583.36,	1931.28
6, 40,	5722.40,	1583.55,	3147.91,	1564.35
6, 40,	7283.29,	1536.41,	3293.64,	1757.23
6, 40,	6777.31,	1592.00,	3124.55,	1532.54
6, 40,	6448.53,	1515.75,	2633.47,	1117.72
6, 40,	7199.34,	1599.09,	3779.62,	2180.53
7, 40,	7458.30,	1718.87,	4657.37,	2938.50
7, 40,	7438.44,	1642.44,	3483.45,	1841.01
7, 40,	7085.85,	1612.77,	3606.59,	1993.82
7, 40,	7472.83,	1630.95,	3900.36,	2269.41
7, 40,	7069.39,	1628.71,	3998.25,	2369.54
8, 40,	6225.19,	1581.77,	2824.99,	1243.22
8, 40,	7447.71,	1680.06,	3804.96,	2124.90
8, 40,	7044.47,	1705.08,	2985.26,	1280.19
8, 40,	7039.96,	1576.66,	3728.38,	2151.72
8, 40,	6151.65,	1637.93,	3219.29,	1581.36
9, 40,	7984.43,	1547.72,	4233.50,	2685.78
9, 40,	8147.43,	1693.16,	4837.27,	3144.11
9, 40,	7462.05,	1592.16,	5099.68,	3507.52
9, 40,	8443.23,	1599.09,	3879.87,	2280.78
9, 40,	7575.24,	1707.51,	3576.77,	1869.26
10, 40,	6375.98,	1432.71,	4216.02,	2783.31
10, 40,	6489.74,	1438.25,	3056.32,	1618.07
10, 40,	6683.16,	1469.11,	3683.84,	2214.73
10, 40,	7133.00,	1437.19,	3274.38,	1837.18
10, 40,	7264.38,	1456.82,	3921.89,	2465.07

# Annexe 6:

DENSITY / HISTOGRAM OVERPLOT WITH MODEL 1  
AND SAMPLE : F5



DENSITIES

D

1.21921E- 3

X-VALUES

X-MIN

1.85092E+ 4

X-MAX

1.99856E+ 4

INTERVAL WIDTH

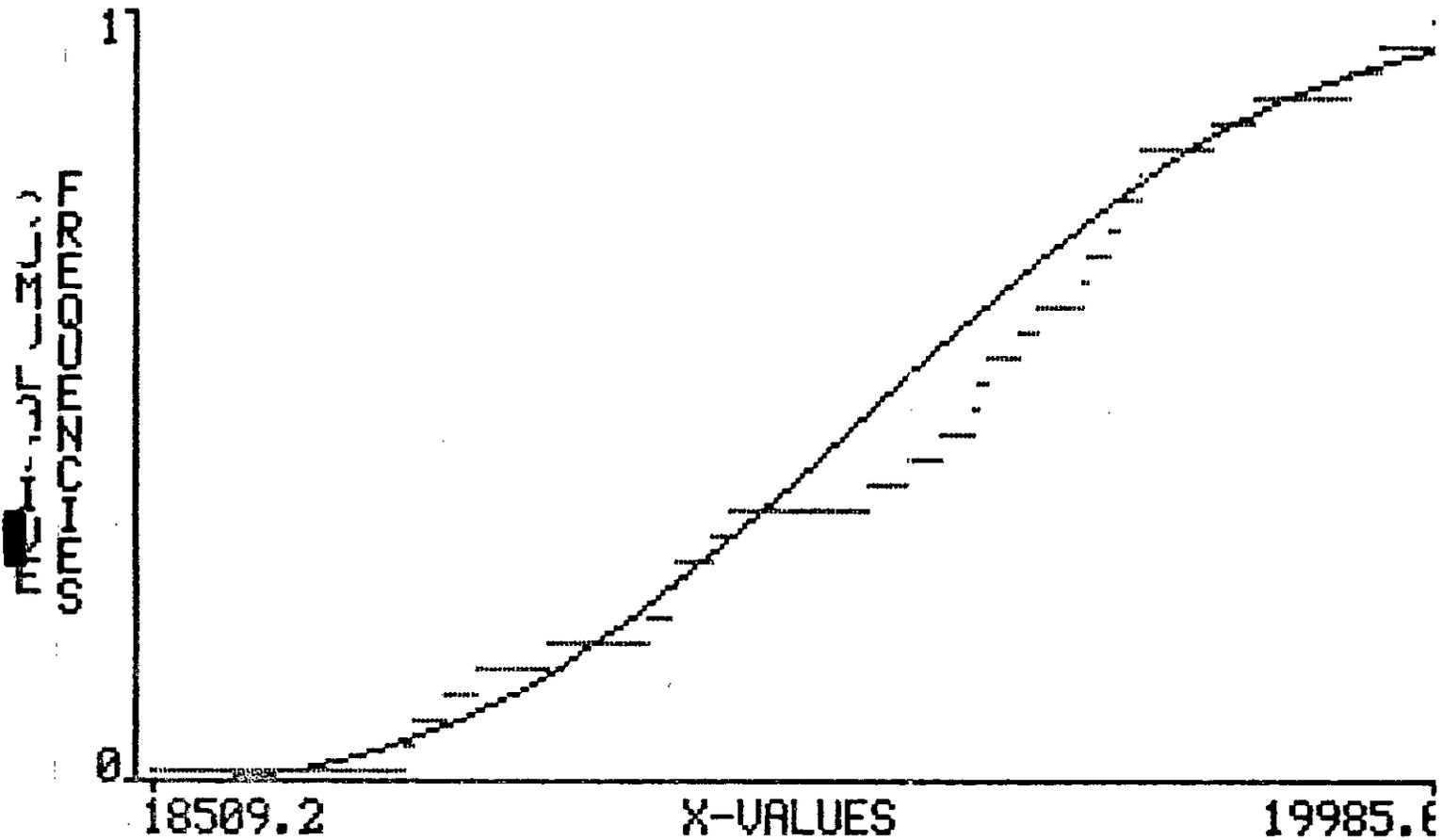
246.060

1-JAN-1980 12:29 A

Densités pour la taille  
égale à 50.  
Résultats du tableau 4.

Annexe 6:  
(suite)

CUMULATIVE FREQUENCY COMPARISON OF MODEL 1  
AND SAMPLE : F5

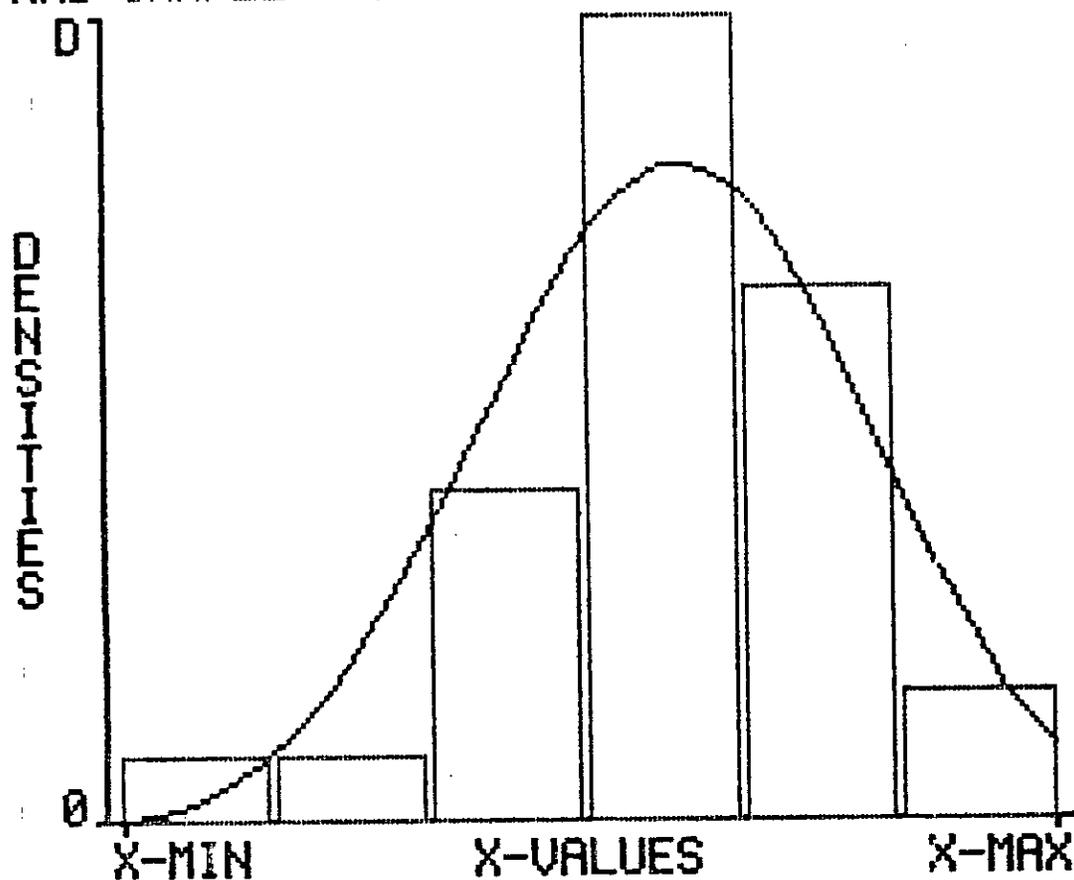


1-JAN-1980 12:32 AI

Ajustement de weibull  
pour la taille du  
problème égale à 50  
(Fréquences cumulées)  
Résultats du tableau 4.

Annexe 6:  
(suite)

DENSITY / HISTOGRAM OVERPLOT WITH MODEL 1  
AND SAMPLE : F8



DENSITIES

D

8.37679E-

X-VALUES

X-MIN

2.38217E+

X-MAX

2.66868E+

INTERVAL WIDT

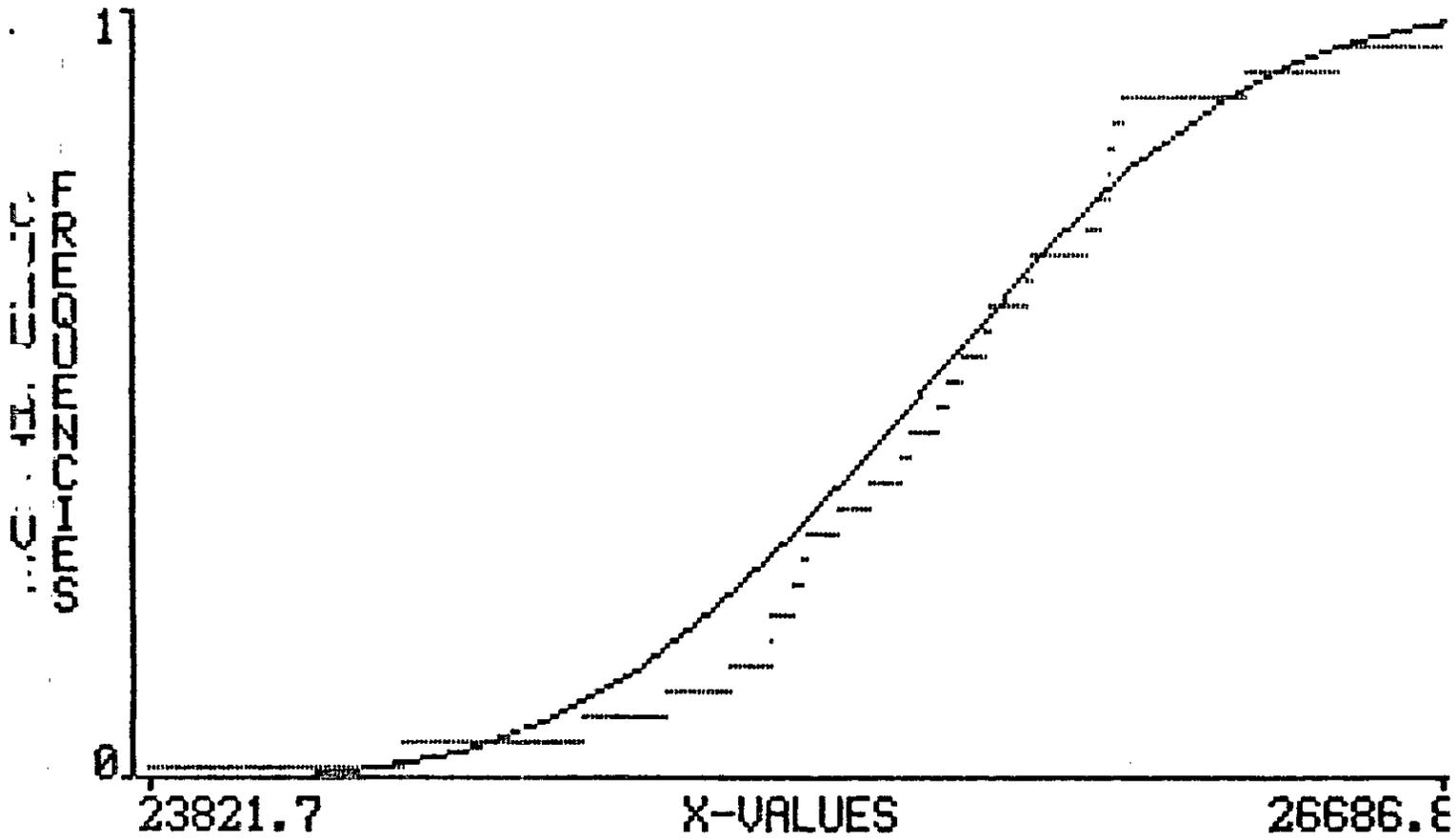
477.510

1-JAN-1980 12:43

Taille 80.

Annexe 6s  
(suite)

CUMULATIVE FREQUENCY COMPARISON OF MODEL 1  
IND SAMPLE : F8

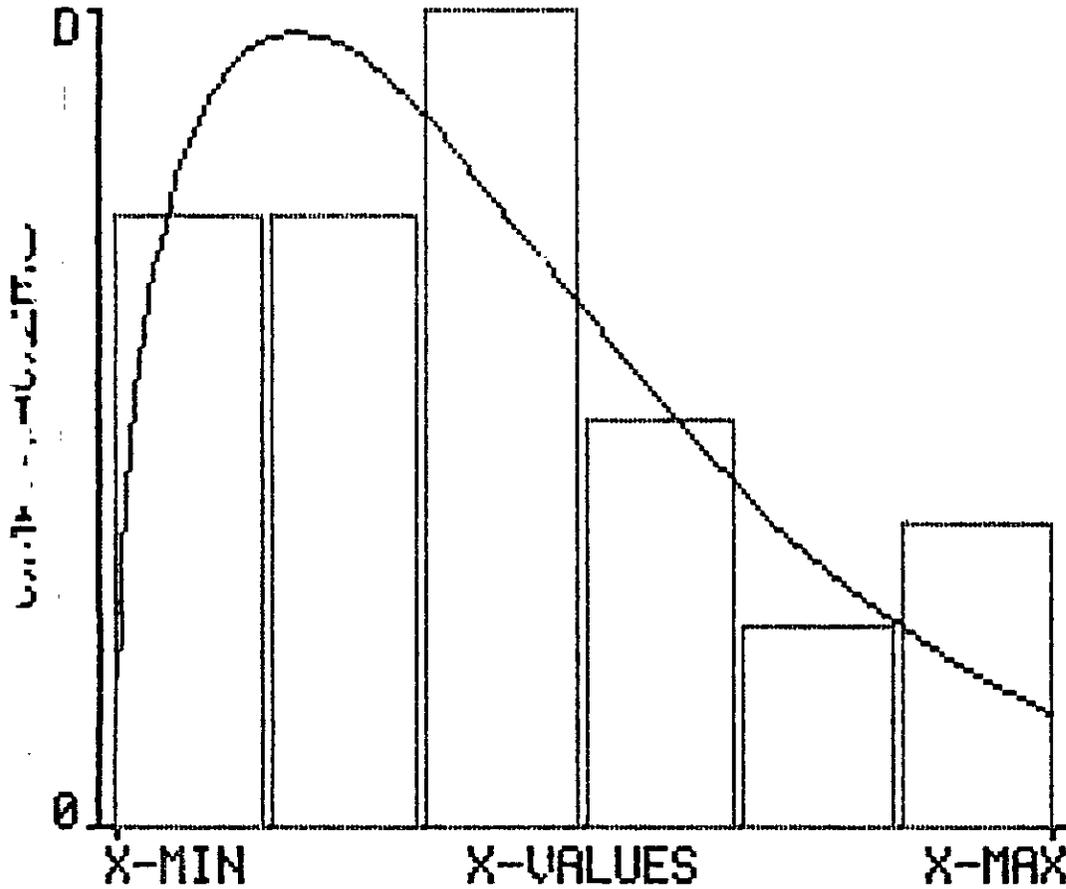


1-JAN-1980 12:45 AM

Taille 80.

Annexe 6:  
(suite)

DENSITY / HISTOGRAM OVERPLOT WITH MODEL 1  
AND SAMPLE : #10



DENSITIES

D

6.74935E- 4

X-VALUES

X-MIN

2.75366E+ 4

X-MAX

2.99072E+ 4

INTERVAL WIDTH

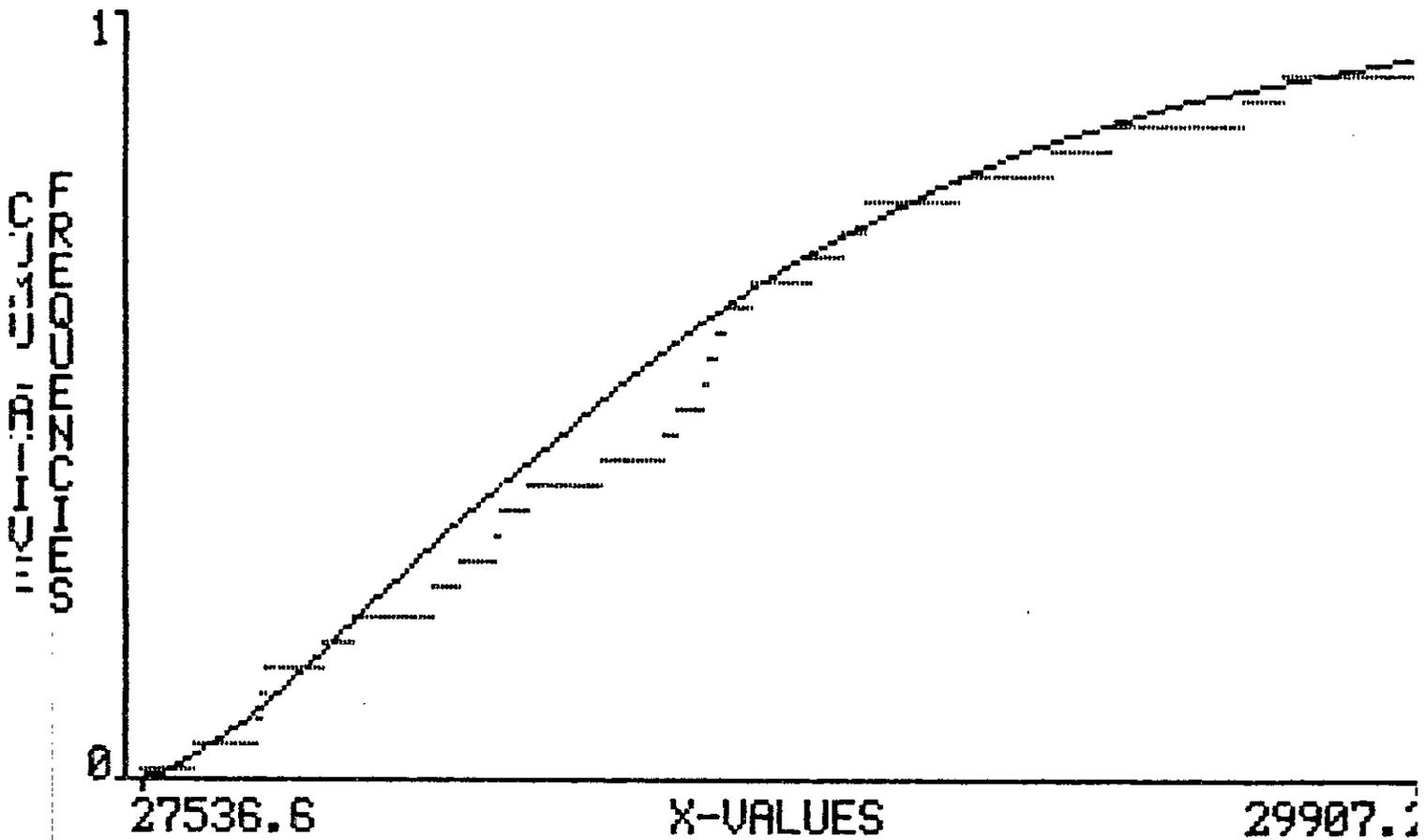
395.100

1-JAN-1980 12:48 AI

Densités pour la taille  
égale à 100.

Annexe 6:  
(suite)

CUMULATIVE FREQUENCY COMPARISON OF MODEL 1  
AND SAMPLE : #10

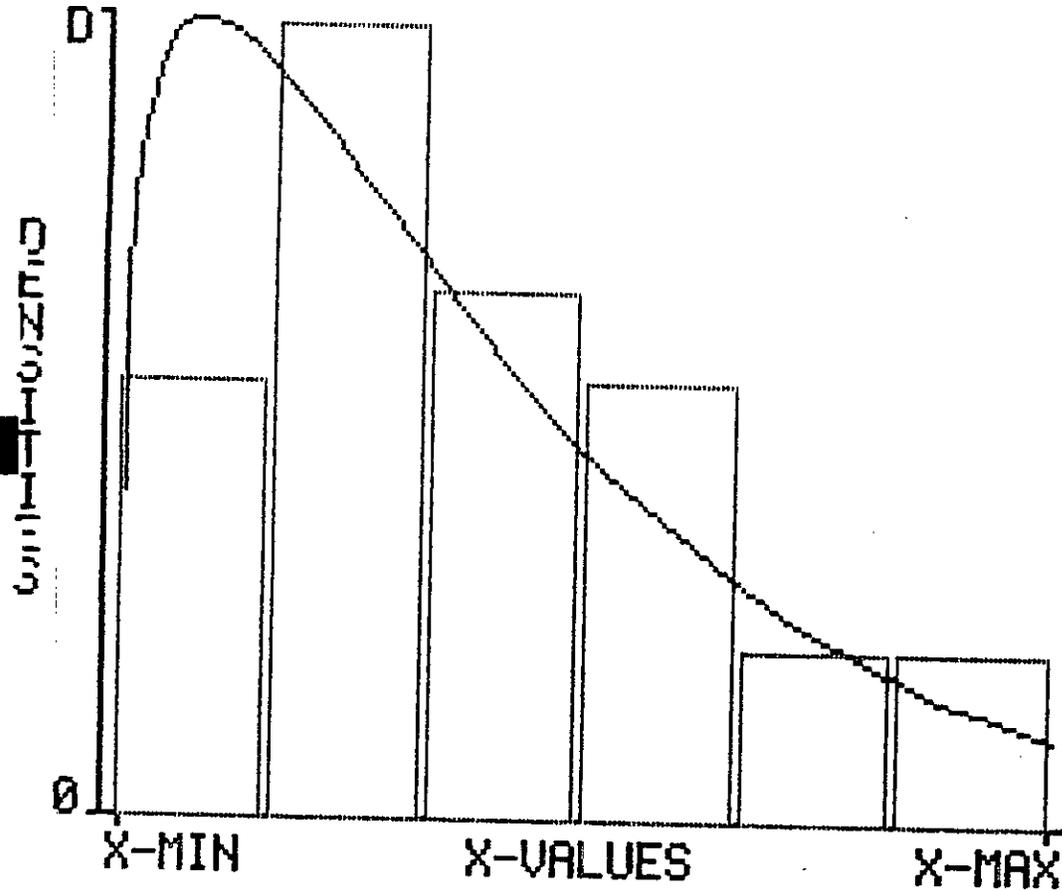


1-JAN-1980 12:52 AI

Ajustement de Weibull  
pour la taille du  
problème égale à 100.  
(Fréquences cumulées)  
Résultats du tableau 4.

Annexe 6:  
(Suite)

DENSITY / HISTOGRAM OVERPLOT WITH MODEL 1  
AND SAMPLE # 04



DENSITIES

D

2.09755E- 3

X-VALUES

X-MIN

1.62869E+ 4

X-MAX

1.71553E+ 4

INTERVAL WIDTH

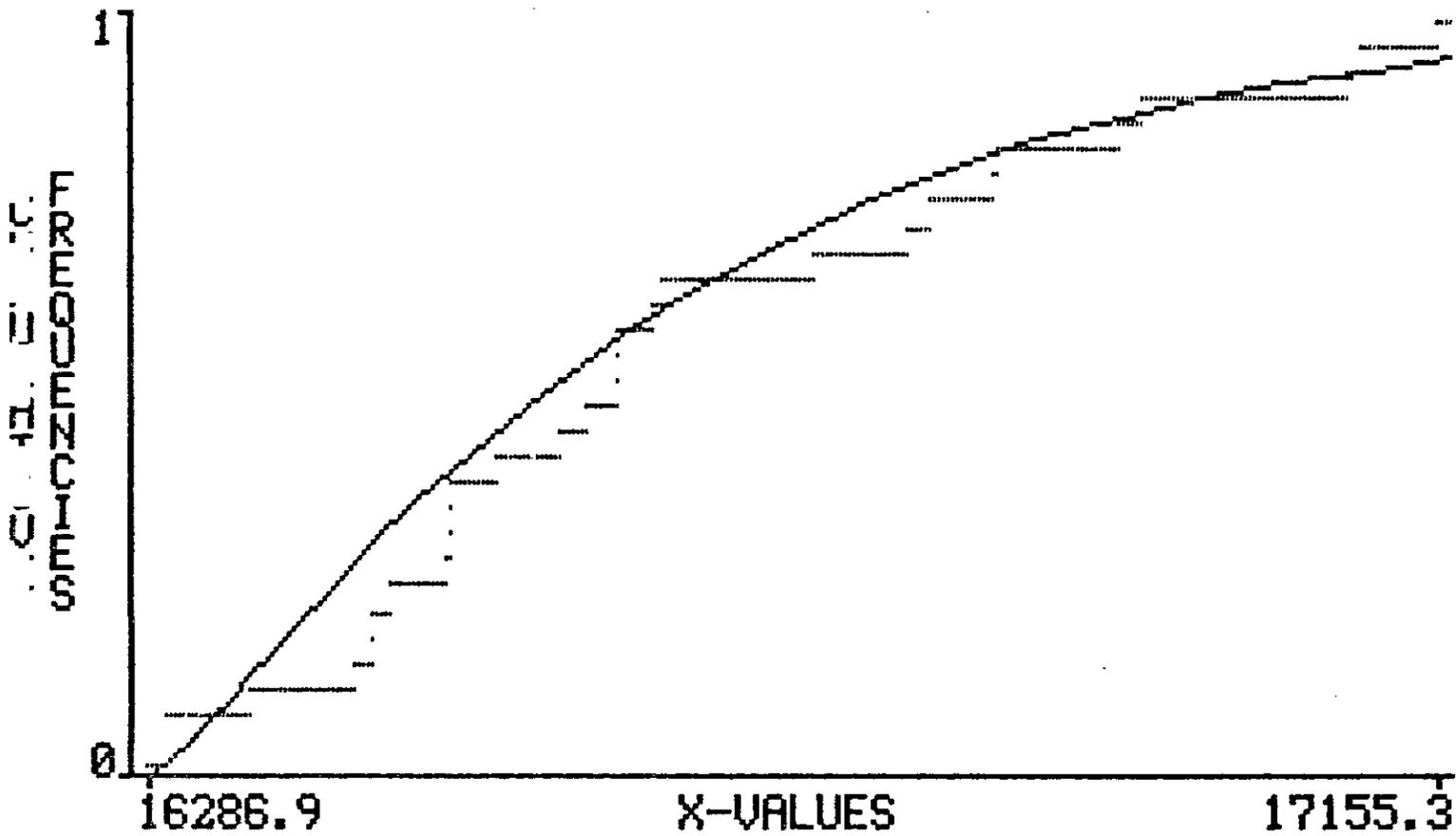
144.730

1-JAN-1980 12:05 AM

Taille 40  
du Tableau 3:

Annexe 6 :  
(suite)

CUMULATIVE FREQUENCY COMPARISON OF MODEL 1  
IND SAMPLE : 04

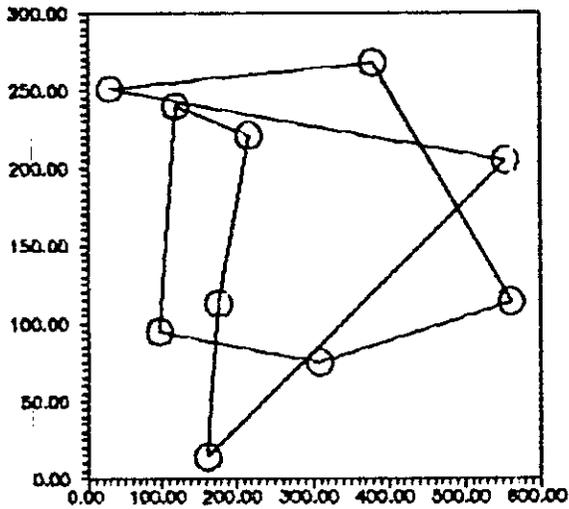


1-JAN-1980 12:09 AM

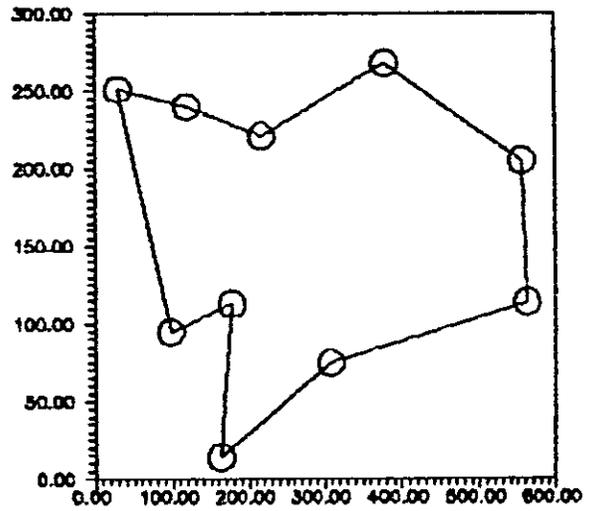
Taille 40.  
du Tableau 3.

# Annexe 7:

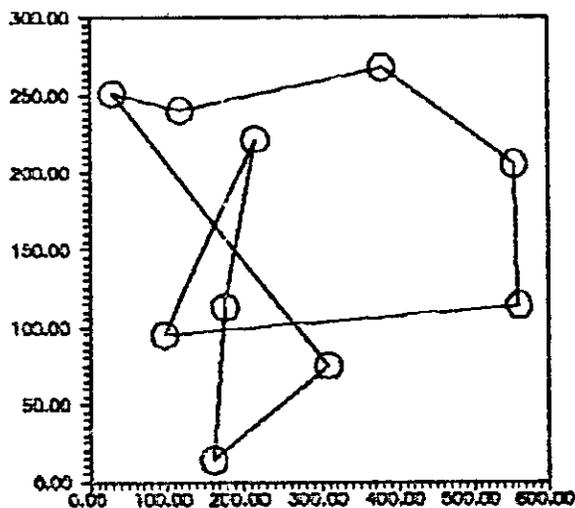
Solution absolue



Solution 2-opt



Solution 3-opt

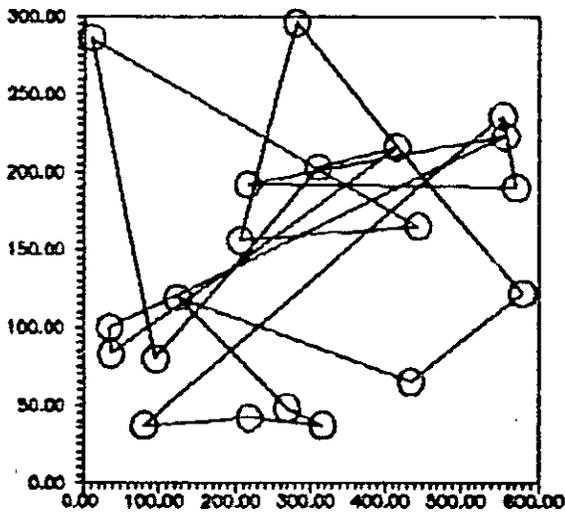


Solutions obtenues par la methode  
2-Opt et 3-Opt a partir d'une tournee  
generee aleatoirement pour la taille 10.

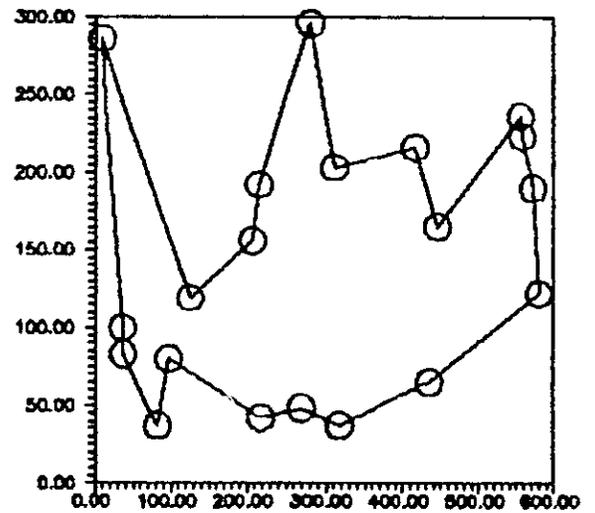
Solutions obtenues par 2-opt  
et 3-opt ; taille 10.

# Annexe 7: (Suite)

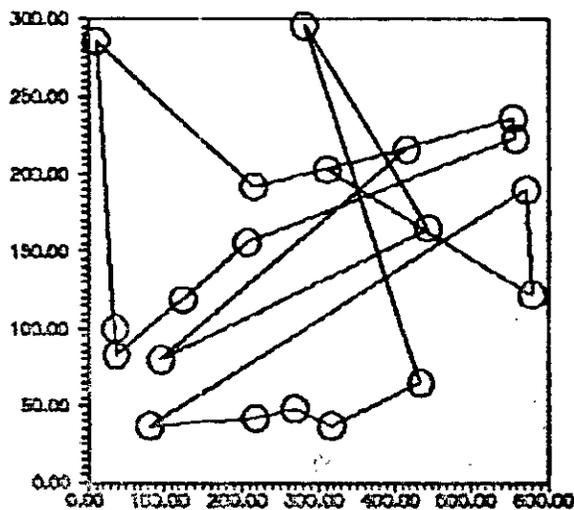
Solution étoilée



Solution 2-opt



Solution 3-opt

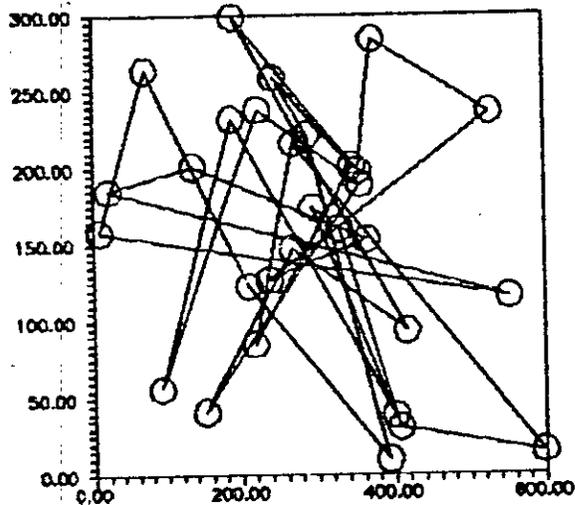


Solutions obtenues par la methode  
2-Opt et 3-Opt a partir d'une tournee  
generee aléatoirement pour la table 20.

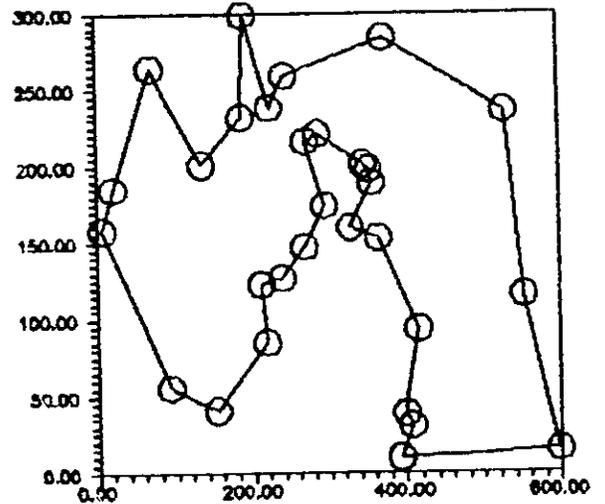
Solutions obtenues par 2-opt  
et 3-opt ; taille 20.

# Annexe 7: (Suite)

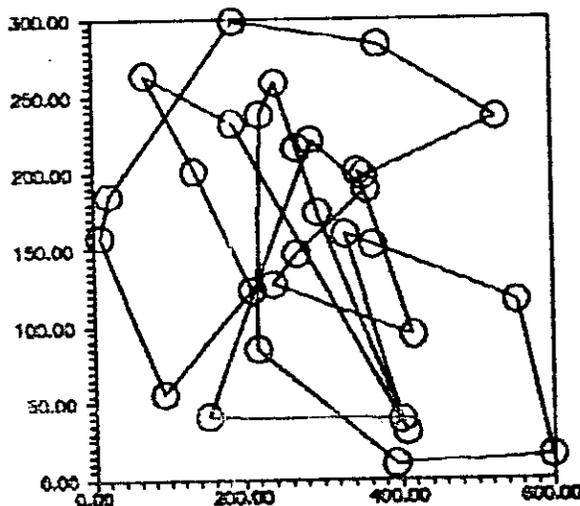
Solution aléatoire



Solution 2-opt



Solution 3-opt



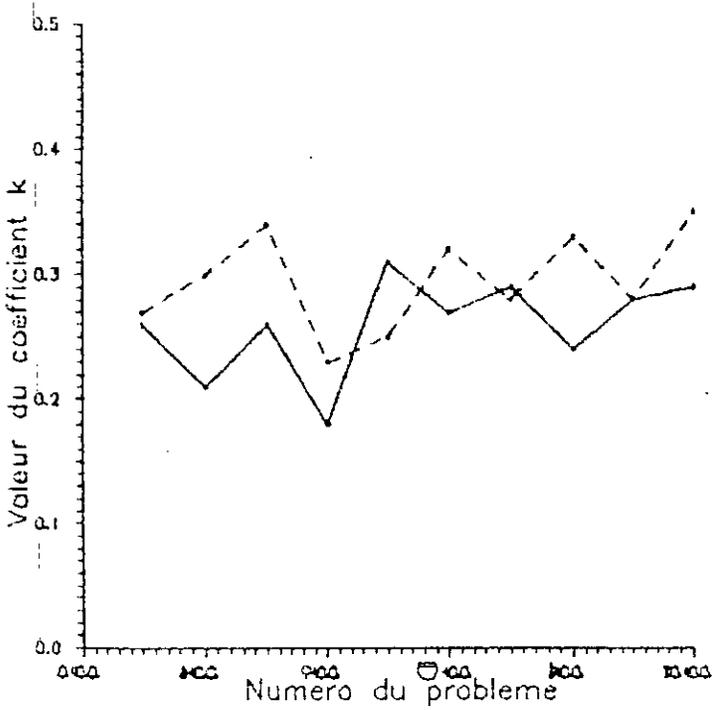
Solutions obtenues par la methode  
2-Opt et 3-Opt a partir d'une tournée  
générée aléatoirement pour la taille 30.

Solutions obtenues par 2-opt  
et 3-opt ; taille 30.

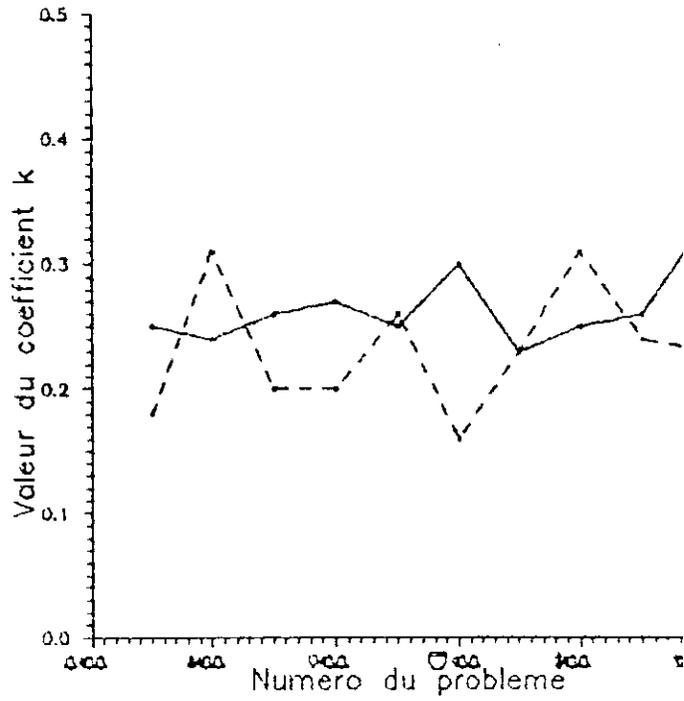
# Annexe 8:

Evolution du coefficient k en fonction de la taille du problème et ceci pour 10 problèmes de taille (15, 20, 25, 30, 35, 40, 50, et 60)

----- taille 15  
 - - - - - taille 20



----- taille 25  
 - - - - - taille 30



# Annexe 8: (suite)

