



HAL
open science

Application d'un bioalgorithme pour le développement d'un procédé de distillation industrielle

Kévin Franiatte

► **To cite this version:**

Kévin Franiatte. Application d'un bioalgorithme pour le développement d'un procédé de distillation industrielle. Sciences pharmaceutiques. 2018. hal-03297586

HAL Id: hal-03297586

<https://hal.univ-lorraine.fr/hal-03297586>

Submitted on 23 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-thesesexercice-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

UNIVERSITE DE LORRAINE 2018

FACULTE DE PHARMACIE

THESE

Présentée et soutenue publiquement

le 27 Septembre 2018 sur un sujet dédié à :

Application d'un bioalgorithme pour le développement d'un procédé de distillation industrielle

pour obtenir

le Diplôme d'Etat de Docteur en Pharmacie

par Kevin FRANIATTE

né le 26 Décembre 1990 à Nancy (54)

Membres du Jury

Président et Co-Directeur :	M. Michel BOISBRUN, MCU, Université de Lorraine
Directeur :	M. Jean-Etienne FORTIER, Pharmacien-ingénieur, UCB Pharma
Membres du jury :	M. Loïc RONIN, Pharmacien-ingénieur, Baxter
	M. Nicolas BALLOIS, Pharmacien-ingénieur, UCB Pharma

UNIVERSITÉ DE LORRAINE
FACULTÉ DE PHARMACIE
Année universitaire 2017-2018

DOYEN

Raphaël DUVAL

Vice-Doyen

Julien PERRIN

Directrice des études

Marie SOCHA

Conseil de la Pédagogie

Présidente, Brigitte LEININGER-MULLER

Vice-Présidente, Alexandrine LAMBERT

Collège d'Enseignement Pharmaceutique Hospitalier

Présidente, Béatrice DEMORE

Commission Prospective Facultaire

Président, Christophe GANTZER

Vice-Président, Jean-Louis MERLIN

Commission de la Recherche

Présidente, Caroline GAUCHER

Chargés de Mission

Communication

Innovation pédagogique

Référente ADE

Référent dotation sur projet (DSP)

Marie-Paule SAUDER

Alexandrine LAMBERT

Virginie PICHON

Dominique DECOLIN

Responsabilités

Filière Officine

Filière Industrie

Filière Hôpital

Pharma Plus ENSIC

Pharma Plus ENSAIA

Pharma Plus ENSGSI

Cellule de Formation Continue et Individuelle

Commission d'agrément des maîtres de stage

ERASMUS

Caroline PERRIN-SARRADO

Julien GRAVOULET

Isabelle LARTAUD

Jean-Bernard REGNOUF de VAINS

Béatrice DEMORE

Marie SOCHA

Jean-Bernard REGNOUF de VAINS

Xavier BELLANGER

Igor CLAROT

Béatrice FAIVRE

François DUPUIS

Mihayl VARBANOV

DOYENS HONORAIRES

Chantal FINANCE

Francine PAULUS

Claude VIGNERON

PROFESSEURS EMERITES

Jeffrey ATKINSON

Jean-Claude BLOCK

Max HENRY

Alain MARSURA

Claude VIGNERON

PROFESSEURS HONORAIRES

Pierre DIXNEUF

Chantal FINANCE

Marie-Madeleine GALTEAU

Thérèse GIRARD

Michel JACQUE

MAITRES DE CONFERENCES HONORAIRES

Monique ALBERT

Mariette BEAUD

François BONNEAUX

Gérald CATAU

Jean-Claude CHEVIN

Pierre LABRUDE
Vincent LOPPINET
Alain NICOLAS
Janine SCHWARTZBROD
Louis SCHWARTZBROD

ASSISTANTS HONORAIRES

Marie-Catherine BERTHE
Annie PAVIS

Jocelyne COLLOMB
Bernard DANGIEN
Marie-CLAUDE FUZELLIER
Françoise HINZELIN
Marie-Hélène LIVERTOUX
Bernard MIGNOT
Blandine MOREAU
Dominique NOTTER
Francine PAULUS
Christine PERDICAKIS
Marie-France Pochon
Anne ROVEL
Gabriel TROCKLE
Maria WELLMAN-ROUSSEAU
Colette ZINUTTI

ENSEIGNANTS

*Section
CNU**

Discipline d'enseignement

PROFESSEURS DES UNIVERSITES-PRATICIENS HOSPITALIERS

Danièle BENSOUSSAN-LEJZEROWICZ	82	<i>Thérapie cellulaire</i>
Jean-Louis MERLIN	82	<i>Biologie cellulaire</i>
Jean-Michel SIMON	81	<i>Economie de la santé, Législation pharmaceutique</i>
Nathalie THILLY	81	<i>Santé publique et Epidémiologie</i>

PROFESSEURS DES UNIVERSITES

Christine CAPDEVILLE-ATKINSON	86	<i>Pharmacologie</i>
Igor CLAROT	85	<i>Chimie analytique</i>
Joël DUCOURNEAU	85	<i>Biophysique, Acoustique, Audioprothèse</i>
Raphaël DUVAL	87	<i>Microbiologie clinique</i>
Béatrice FAIVRE	87	<i>Hématologie, Biologie cellulaire</i>
Luc FERRARI	86	<i>Toxicologie</i>
Pascale FRIANT-MICHEL	85	<i>Mathématiques, Physique</i>
Christophe GANTZER	87	<i>Microbiologie</i>
Frédéric JORAND	87	<i>Eau, Santé, Environnement</i>
Isabelle LARTAUD	86	<i>Pharmacologie</i>
Dominique LAURAIN-MATTAR	86	<i>Pharmacognosie</i>
Brigitte LEININGER-MULLER	87	<i>Biochimie</i>
Pierre LEROY	85	<i>Chimie physique</i>
Philippe MAINCENT	85	<i>Pharmacie galénique</i>
Patrick MENU	86	<i>Physiologie</i>
Jean-Bernard REGNOUF de VAINS	86	<i>Chimie thérapeutique</i>
Bertrand RIHN	87	<i>Biochimie, Biologie moléculaire</i>

MAITRES DE CONFERENCES DES UNIVERSITES-PRATICIENS HOSPITALIERS

Béatrice DEMORE	81	<i>Pharmacie clinique</i>
Alexandre HARLE	82	<i>Biologie cellulaire oncologique</i>
Julien PERRIN	82	<i>Hématologie biologique</i>
Loïc REPEL	82	<i>Biothérapie</i>
Marie SOCHA	81	<i>Pharmacie clinique, thérapeutique et biotechnique</i>

MAITRES DE CONFERENCES

Sandrine BANAS	87	<i>Parasitologie</i>
Xavier BELLANGER	87	<i>Parasitologie, Mycologie médicale</i>
Emmanuelle BENOIT	86	<i>Communication et Santé</i>
Isabelle BERTRAND	87	<i>Microbiologie</i>
Michel BOISBRUN	86	<i>Chimie thérapeutique</i>
Ariane BOUDIER	85	<i>Chimie Physique</i>
Cédric BOURA	86	<i>Physiologie</i>

Joël COULON	87	Biochimie
Sébastien DADE	85	Bio-informatique
Dominique DECOLIN	85	Chimie analytique
Roudayna DIAB	85	Pharmacie galénique
Natacha DREUMONT	87	Biochimie générale, Biochimie clinique
Florence DUMARCAZ	86	Chimie thérapeutique
François DUPUIS	86	Pharmacologie
Reine EL OMAR	86	Physiologie
Adil FAIZ	85	Biophysique, Acoustique
Anthony GANDIN	87	Mycologie, Botanique
Caroline GAUCHER	86	Chimie physique, Pharmacologie
Stéphane GIBAUD	86	Pharmacie clinique
Thierry HUMBERT	86	Chimie organique
Olivier JOUBERT	86	Toxicologie, Sécurité sanitaire
Alexandrine LAMBERT	85	Informatique, Biostatistiques
Julie LEONHARD	86/01	Droit en Santé
Christophe MERLIN	87	Microbiologie environnementale
Maxime MOURER	86	Chimie organique
Coumba NDIAYE	86	Epidémiologie et Santé publique
Marianne PARENT	85	Pharmacie galénique
Caroline PERRIN-SARRADO	86	Pharmacologie
Virginie PICHON	85	Biophysique
Sophie PINEL	85	Informatique en Santé (e-santé)
Anne SAPIN-MINET	85	Pharmacie galénique
Marie-Paule SAUDER	87	Mycologie, Botanique
Guillaume SAUTREY	85	Chimie analytique
Rosella SPINA	86	Pharmacognosie
Sabrina TOUCHET	86	Pharmacochimie
Mihayl VARBANOV	87	Immuno-Virologie
Marie-Noëlle VAULTIER	87	Mycologie, Botanique
Emilie VELOT	86	Physiologie-Physiopathologie humaines
Mohamed ZAIUO	87	Biochimie et Biologie moléculaire

PROFESSEURS ASSOCIES

Julien GRAVOULET	86	Pharmacie clinique
Anne MAHEUT-BOSSER	86	Sémiologie

PROFESSEUR AGREGE

Christophe COCHAUD	11	Anglais
--------------------	----	---------

*Disciplines du Conseil National des Universités :

80 : Personnels enseignants et hospitaliers de pharmacie en sciences physico-chimiques et ingénierie appliquée à la santé

81 : Personnels enseignants et hospitaliers de pharmacie en sciences du médicament et des autres produits de santé

82 : Personnels enseignants et hospitaliers de pharmacie en sciences biologiques, fondamentales et cliniques

85 : Personnels enseignants-chercheurs de pharmacie en sciences physico-chimiques et ingénierie appliquée à la santé

86 : Personnels enseignants-chercheurs de pharmacie en sciences du médicament et des autres produits de santé

87 : Personnels enseignants-chercheurs de pharmacie en sciences biologiques, fondamentales et cliniques

11 : Professeur agrégé de lettres et sciences humaines en langues et littératures anglaises et anglo-saxonnes

SERMENT DES APOTHICAIRES



Je jure, en présence des maîtres de la Faculté, des conseillers de l'ordre des pharmaciens et de mes condisciples :

D' honorer ceux qui m'ont instruit dans les préceptes de mon art et de leur témoigner ma reconnaissance en restant fidèle à leur enseignement.

D' exercer, dans l'intérêt de la santé publique, ma profession avec conscience et de respecter non seulement la législation en vigueur, mais aussi les règles de l'honneur, de la probité et du désintéressement.

D' e ne jamais oublier ma responsabilité et mes devoirs envers le malade et sa dignité humaine ; en aucun cas, je ne consentirai à utiliser mes connaissances et mon état pour corrompre les mœurs et favoriser des actes criminels.

Que les hommes m'accordent leur estime si je suis fidèle à mes promesses.

Que je sois couvert d'opprobre et méprisé de mes confrères si j'y manque.

« LA FACULTE N'ENTEND DONNER AUCUNE APPROBATION, NI
IMPROBATION AUX OPINIONS EMISES DANS LES THESES, CES
OPINIONS DOIVENT ETRE CONSIDEREES COMME PROPRES A
LEUR AUTEUR ».

Remerciements

A Monsieur Boisbrun

Maître de conférences,

Pour m'avoir fait l'honneur d'accepter la présidence de ce jury de thèse, pour votre bienveillance et votre disponibilité malgré la distance.

Veillez trouver ici le témoignage de ma profonde considération.

A Monsieur Fortier

Pharmacien-ingénieur ENSIC,

Pour m'avoir fait l'honneur de diriger cette thèse, pour ta disponibilité, tes conseils avisés et toutes ces discussions stimulantes au sujet des procédés industriels. Je tiens également à te remercier de m'avoir accepté en tant que stagiaire dans ton département et de m'avoir toujours soutenu depuis toutes ces années.

Je t'adresse ici toute ma gratitude et l'expression de tout mon respect.

A mes juges,

Monsieur Ronin

Pharmacien-ingénieur ENSIC,

Pour ton aide très détaillée, ta disponibilité, ta motivation et tes précieux conseils.

Je t'adresse ici mes plus vifs remerciements.

Monsieur Ballois

Pharmacien-ingénieur ENSAIA,

Pour tout l'intérêt que tu as porté à ce sujet, les échanges constructifs associés, ta participation dans ce jury et ta disponibilité en cette période de changement professionnel.

Je t'adresse ici l'expression de ma profonde reconnaissance.

A ma famille,

A toi **maman**,

Qui vit les moments importants de ma vie avec encore plus d'intensité que moi. Merci pour ta présence et ton soutien sans limites qui m'ont rendu la vie plus agréable chaque jour depuis toutes ces années.

A toi **papa**

Pour ces conseils avisés qui me poussent toujours à vouloir faire mieux, pour m'avoir encouragé sans relâche à écrire ce document ("Et ta thèse ?") et pour tout ce que tu m'as enseigné.

A **Lara**,

Ma coéquipière dans la vie, pour m'avoir épaulé moralement tous les jours dans la construction de cette

thèse.

Ma réussite est également la vôtre. Je vous adresse toute ma gratitude du fond du coeur.

A mes amis,

Mes fidèles compagnons : Samy, David, Rosyne, Arthur, Matthieu, Sylvain, Alban, Adrien, Julia, Sophie, Magali, Alex, Manon, Lise, Matthew, Jérôme, Christophe et Morgane.

La clique de la faculté : Marc, Edouard, Charlotte, Saber, Alexandre, Loïc, Emmanuelle et Victoria.

La bande de l'ENSIC : Mathieu, Charles, Jérémy, Sébastien, Aurélien, Tiphaine et Julia.

A Caroline, tout particulièrement, qui a su me motiver pour écrire cette thèse.

Et tous ceux avec qui j'ai partagé de si bons moments durant toutes ces années.

En témoignage de mon affection.

Musique écoutée lors de l'écriture de ce document : Meshuggah et Jazz Bossa Nova.

Table des matières

Liste des abréviations	7
Introduction	9
1 Les algorithmes biologiques	11
1.1 Les bio-algorithmes et les systèmes évolutifs	12
1.1.1 L'optimisation en essaim de particules ou Particle Swarm Optimization (PSO)	12
1.1.2 Les algorithmes inspirés des colonies de fourmis ou Ant Colony Optimization (ACO)	12
1.1.3 Les algorithmes basés sur les colonies d'abeilles ou Artificial Bee Colony Optimization (ABC)	13
1.1.4 L'évolution différentielle	14
1.1.5 Les algorithmes bactériens	14
1.2 Comparaison avec les méthodes classiques	15
1.3 L'algorithme Krill Herd	16
1.3.1 Modèle Lagrangien	16
1.3.2 Mouvement induit par les autres krills	17
1.3.3 Recherche de nourriture	18
1.3.4 Diffusion physique	20
1.3.5 Modèle itératif de l'algorithme	20
1.3.6 Les opérateurs génétiques	21
1.3.7 Critère d'arrêt, méthodologie et programme	24
1.3.8 Fonctions d'évaluation de la performance de l'algorithme	24
2 Construction d'un diagramme de phases à l'aide d'un bioalgorithme	29
2.1 Calcul de flash par minimisation de l'énergie de Gibbs de mélange	29
2.1.1 Première et deuxième loi de la thermodynamique	29
2.1.2 Les conditions d'équilibre entre phases	30
2.1.3 Energie libre de Gibbs	30
2.1.4 Grandeurs de mélange et solutions réelles liquides	32
2.1.5 Activité	33
2.1.6 Fugacité	34
2.2 Calcul des équilibres liquide-vapeur du binaire méthyl tert-butyl-éther / eau	34
2.3 Calcul des pressions de vapeur saturante	37
2.4 Modèle NRTL	38
2.5 Construction du diagramme des équilibres liquide-vapeur	42
2.5.1 Types de diagrammes d'équilibres liquide - vapeur	42
2.5.2 Résultats pour le système méthyl tert-butyl-éther/eau	43

3	Application à un cas concret de distillation industrielle	47
3.1	Objectifs de la distillation industrielle	47
3.1.1	Distillation flash	47
3.1.2	Distillation simple	49
3.1.3	Distillation fractionnée	49
3.2	Cas pratique : distillation à échelle pilote dans la dernière étape de synthèse d'un principe actif	50
3.2.1	Objectifs	52
3.2.2	Distillation simple batch	53
3.2.3	Distillation simple continue	57
3.2.4	Résultats de la campagne à échelle pilote	61
3.3	Pistes d'amélioration	63
3.3.1	Amélioration du modèle	63
3.3.2	Amélioration du procédé	64
	Conclusion	67
	Bibliographie	96

Liste des illustrations

1.1	Figure représentant le jeu de Pentamino, à gauche les pièces du puzzle, à droite la figure à représenter [10]	13
1.2	Schéma représentatif du principe de la distance de détection appliqué à chaque krill pour déterminer son voisinage [29]	18
1.3	Schéma représentatif de l'effet de β_i , d'un côté la composante β_i^{best} qui attire le krill i vers sa meilleure position visitée, et β_i^{food} l'effet de centre de masse	20
1.4	Schéma illustrant un exemple de situation d'intérêt des opérateurs génétiques, en bleu l'individu muté ou croisé, en vert l'individu ayant la meilleure valeur avant mutation ou croisement	23
1.5	Graphiques des fonctions d'Ackley (en haut) et de Rastrigin (en bas) pour $n = 2$, le minimum global des deux fonctions est 0 [30, 31]	25
1.6	Graphique semi-logarithmique des valeurs renvoyées par l'algorithme Krill Herd au fur et à mesure des itérations pour la fonction d'Ackley (en haut) et de Rastrigin (en bas), le nombre de dimensions est fixé à $n = 20$	26
2.1	Energie libre de Gibbs molaire G d'un système binaire stable en fonction de la fraction molaire en composé B	31
2.2	Variation de l'énergie libre de Gibbs molaire G d'un système binaire en fonction de la fraction molaire en composé B	32
2.3	Courbe de vapeur saturante du méthyl-ter-butyl-ether entre 0°C et 80°C (en rouge) et de l'eau entre 0°C et 100°C (en bleu) donnés par l'équation d'Antoine	38
2.4	Arbre décisionnel pour le choix du modèle à employer selon la solution réelle considérée. En rouge le chemin suivi pour l'exemple méthyl tert-butyl-éther - eau [50]	39
2.5	Schéma représentatif des interactions prises en compte dans le modèle NRTL d'un mélange binaire (molécules 1 et 2)	40
2.6	Schémas des différents diagrammes liquide-vapeur usuels rencontrés pour des systèmes binaires à pression constante en fonction de la fraction molaire ou massique de l'un des deux constituants et de la température [53]	42
2.7	Diagramme (x,T) des équilibres liquide-vapeur du mélange méthyl tert-butyl-éther/eau à pression atmosphérique généré par l'algorithme Krill Herd, en bleu la composition liquide, en orange la composition vapeur	44
2.8	Energie libre de Gibbs de mélange du système méthyl tert-butyl-éther/eau en fonction de la fraction molaire en eau à 52.04°C et à pression atmosphérique	45
3.1	Schéma exemple d'un diagramme des équilibres liquide-vapeur d'un binaire A et B , la température T est fonction de la composition en composé A dans le mélange [59]	48
3.2	Schéma général du procédé de la distillation flash pour deux cas : entrée liquide ou vapeur [60]	48
3.3	Schéma d'une distillation fractionnée avec une colonne à plateaux [64]	50
3.4	Diagramme explicatif du procédé de la dernière étape de synthèse du principe actif	51
3.5	Diagramme de phase méthyl tert-butyl-éther/eau à pression atmosphérique	52

3.6	Schéma d'un exemple d'une distillation de type batch	55
3.7	Schéma de la distillation de type batch réalisée à échelle pilote à UCB Pharma	55
3.8	Graphique de $\frac{1}{y-x}$ en fonction de la fraction molaire en eau x_R avec l'aire sous la courbe considérée pour la distillation batch	56
3.9	Schéma d'une distillation de type continue	57
3.10	Schéma de la distillation de type continu proposée pour l'échelle pilote	59
3.11	Graphique de $\frac{1}{y}$ en fonction de la fraction molaire en eau x_R avec l'aire sous la courbe considérée pour la distillation continue	59
3.12	Graphique comparatif de la fraction molaire en eau dans la phase liquide en fonction de la quantité molaire de distillats collectés pour la distillation continue (en orange) et la distillation batch (en bleu) selon les estimations théoriques	60
3.13	Dessin technique du réacteur Pfaudler AE400 utilisé pour la distillation, en rouge la partie couverte par la double enveloppe, en bleu la partie "froide" du réacteur	62
3.14	Diagramme ternaire vierge des équilibres liquide-vapeur pour trois composés A, B et C .	64
3.15	Diagramme binaire des équilibres liquide-vapeur pour l'eau et le méthyl tert-butyl-éther à 0.5, 1.0 et 2.0 bar générés par l'algorithme Krill Herd	65

Liste des tableaux

2.1	Coefficients d'Antoine pour le méthyl-ter-butyl-éther et l'eau, le domaine de validité des coefficients est défini entre T_{min} et T_{max} [46]	37
3.1	Compositions massiques à l'état initial et final pour la distillation du principe actif	53
3.2	Compositions molaires à l'état initial et final pour la distillation du principe actif sans compter la présence du soluté	53
3.3	Compositions molaires finales théoriques pour la distillation batch du principe actif et sans compter la présence du soluté	56
3.4	Compositions molaires finales théoriques pour la distillation continue du principe actif et sans compter la présence du soluté	60
3.5	Compositions molaires finales, sans tenir compte de la présence du soluté, après distillation pour deux batchs de principe actif réalisés à l'échelle pilote à UCB Pharma	61

Liste des abréviations

ABC	Artificial Bee Colony algorithm
ACO	Ant Colony Optimization algorithm
DDB	Dortmund Data Bank
MTBE	Méthyl Tert-Butyl-Ether
NRTL	Non Random Two Liquid
LLE	Liquid-Liquid Equilibrium
PRSK	Predictive Soave-Redlich-Kwong
PSO	Particle Swarm Optimization
QbD	Quality by Design
UCB	Union Chimique Belge
UNIFAC	UNIversal Functionnal Activity Coefficient
UNIQUAC	UNIversal QUAsiChemical
VLE	Vapor-Liquid Equilibrium

Introduction

DANS un contexte de forte compétitivité et de règles de plus en plus strictes, l'industrie pharmaceutique fait face à des défis qui évoluent continuellement. Pour maintenir sa performance et garantir le développement de produits de qualité en un minimum de temps, de coûts et de ressources, il est nécessaire d'employer des stratégies de modernisation de ses approches. C'est pourquoi depuis quelques années, l'ingénierie des procédés pharmaceutiques profite de l'essor de nouveaux outils pour la conception de ses procédés.

La complexité des structures moléculaires des principes actifs est grandissante et les coûts de développement doivent être réduits malgré le nombre considérable d'étapes de synthèse. Chacune de ces étapes comporte des opérations de traitement pour la purification et l'isolement des intermédiaires réactionnels jusqu'au principe actif. Le but est de porter ces procédés à grande échelle de façon robuste et en un minimum de temps tout en limitant le travail expérimental qui peut être long et coûteux.

Plus généralement, le développement pharmaceutique est atypique car il doit alimenter les études cliniques quel que soit le stade de développement du médicament. Les procédés ne sont donc pas optimisés et sont soumis à des risques qualité. Historiquement, l'optimisation comprenait beaucoup d'étapes itératives entre le laboratoire et les essais pilotes nécessaires à la compréhension de l'effet d'échelle avant le transfert en production. Les synthèses de l'ordre du kilogramme sont particulièrement coûteuses et laissent de moins en moins place à l'erreur, d'où l'importance de mettre en place des modèles prédictifs, qui, sur base des observations au laboratoire, permettent de s'affranchir d'un trop grand nombre de "tests" à grande échelle. Du principe actif au produit fini, la modélisation peut être appliquée à la plupart des étapes de la conception d'un médicament et doit être employée le plus tôt possible dans le travail de développement afin d'apporter une amélioration continue durant toute la vie de ce dernier.

La compréhension théorique des phénomènes et des mécanismes en jeu est indispensable ainsi que l'utilisation de technologies innovantes. Cette démarche est en partie initiée par la qualité par la conception (QbD) qui suggère de collecter plus d'informations sur les procédés de fabrication afin de comprendre les processus en jeu, les rendre plus flexibles et mieux gérer les risques associés. Une fois les modèles établis, il est aisé de les adapter pour un large éventail de problématiques et étudier différentes stratégies. Les essais au laboratoire sont donc utiles, au mieux, pour la validation des résultats apportés par ces prédictions.

Dans cette thèse, l'objectif est de présenter un exemple d'outil entièrement développé pour la caractérisation d'un procédé existant afin d'établir une procédure adaptée pour son amélioration. Plus concrètement, les solutions étudiées visent à estimer la consommation minimale en solvant utile à une opération de distillation intervenant dans la dernière étape de synthèse d'un principe actif à usage anti-épileptique actuellement en phase 3 et d'évaluer d'autres pistes pour réduire la consommation des solvants utilisés. Cette étape, plus importante que les précédentes, déterminera également la pureté et la forme des cristaux finaux isolés. Le développement de celle-ci est réalisée au département Recherche et Développement à UCB Pharma en Belgique et comprend au total douze étapes.

Tout d'abord, le concept de bioalgorithme sera présenté. Des bioalgorithmes inspirés de la nature et assimilés à de l'intelligence artificielle, permettent la résolution de problèmes mathématiques relativement complexes sur base de l'observation du comportement d'organismes vivants. Après avoir introduit les plus usuels d'entre eux, un bioalgorithme sera étudié plus particulièrement dans l'objectif de l'implémenter dans un programme de minimisation numérique. Dans une seconde partie, la fonction que l'on souhaite minimiser par le programme sera établie grâce à des notions de thermodynamique. Les concepts théoriques liés à cette fonction seront largement détaillés tout comme le choix des modèles. Au final, un diagramme de phases caractéristique du système de solvants intervenant dans l'opération de distillation considérée pourra être généré. L'ensemble de ces travaux se destine à une utilisation libre du programme proposé et offre une alternative économique aux licences commerciales. Enfin, dans une troisième partie, le diagramme ainsi créé permettra d'adopter la stratégie adéquate pour atteindre les spécifications imposées et minimiser la consommation de solvant frais. On y développera les différents modes et types de distillations industrielles, l'implémentation d'une procédure à l'échelle pilote et les résultats associés.

1 Les algorithmes biologiques

Les algorithmes inspirés de la biologie font partie d'un domaine de recherche devenu important pour la résolution de problèmes complexes et réels [1]. On les retrouve en physique, mais également en économie ou dans les sciences sociales. En imitant les phénomènes rencontrés dans la nature, ces méthodes heuristiques sont maintenant bien reconnues suite aux multiples succès générés [2, 3].

Beaucoup de ces recherches se concentrent sur deux grands axes : l'un dédié à la résolution de problèmes académiques (complexes et difficilement applicables), l'autre pour le développement de nouveaux algorithmes dérivés. Ces travaux ont pour intention de pallier certaines difficultés bien connues comme la convergence lente vers l'optimum, la stationnarité sur un minimum local (et non global) ou encore le temps de calcul élevé. En effet, l'objectif de l'optimisation, en mathématiques, consiste à rechercher un extremum global d'une fonction objectif sur tout son domaine de définition (minimum ou maximum). Les applications sont très variées ; on peut citer par exemple la reconnaissance de motifs, la reconstruction d'objets, le contrôle robotique ou bien l'identification intelligente.

Dès le début de l'ère informatique, il fut essentiel d'optimiser les programmes de résolution pour des raisons de ressources et de temps de calcul. Avec la popularité grandissante des ordinateurs et le volume exponentiel de données, les tâches sont devenues de plus en plus complexes et les algorithmes conventionnels de moins en moins efficaces, il était donc nécessaire d'inclure une certaine intelligence à ces algorithmes. Cette nouvelle forme a donc été nommée intelligence artificielle.

Afin de définir l'intelligence artificielle, il est important de noter que ses nombreuses caractéristiques ne sont pas indépendantes mais combinées de façon synergique. On y retrouvera alors typiquement des éléments d'adaptabilité, de communication, de rétro-contrôle, d'exploration et d'exploitation ainsi qu'une partie aléatoire (ou plus exactement, pseudo-aléatoire). Aucun de ces éléments n'est indépendamment satisfaisant, il y a donc une notion de coopération et d'évolution plutôt que de compétition.

Plusieurs termes sont associés à ces algorithmes biologiques, les plus représentatifs sont :

- le calcul évolutif
- l'optimisation basée sur une population
- le calcul souple ou "soft computing"
- l'apprentissage machine
- l'intelligence collective

Les phénomènes étudiés sont donc très proches de la théorie de la sélection naturelle introduite par Charles Darwin en 1859, qui repose sur quatre piliers : la population, la diversité, l'hérédité et la sélection. L'algorithmique biologique est encore une science jeune qui n'a pas atteint sa maturité, et qui compte beaucoup sur la biologie pour décrire au mieux les comportements des organismes vivants ainsi que le fonctionnement de leur interconnectivité.

1.1 Les bio-algorithmes et les systèmes évolutifs

Le calcul évolutif est une branche de l'intelligence artificielle qui inclut une variété de techniques pour résoudre des problèmes. Il est basé sur les principes de l'évolution biologique et existe depuis les années cinquante. Dans un tel algorithme, chaque individu de la population est assimilé à une solution possible. Les candidats les plus qualifiés sont ceux qui obtiennent les meilleurs résultats sur la fonction objectif. L'évolution de la population est obtenue par itérations successives dans lesquelles une série d'opérations sont appliquées comme la reproduction, la mutation, la recombinaison ou la sélection. Une nouvelle "génération" de la population est alors créée avec de meilleurs résultats potentiels. La manière dont la population évolue est intrinsèque à la méthode utilisée.

1.1.1 L'optimisation en essaim de particules ou Particle Swarm Optimization (PSO)

L'algorithme en essaim de particules, littéralement, a été introduit pour la première fois par Kennedy et Eberhart en 1995. Il est basé sur le comportement social des oiseaux migrateurs [4]. Il exploite une population, ou un essaim, de N solutions potentielles où chaque solution est représentée par un individu i à une itération t . À l'itération suivante $t + 1$, les nouvelles positions de chaque individu sont recalculées, à l'aide d'un mouvement v appliqué à ceux-ci. Il est alors facile de les comparer aux anciennes positions pour déduire, pour chaque individu, si sa position s'est améliorée. L'équation (1.1) représente plus formellement cette séquence itérative. Il faut noter que si la fonction à optimiser comporte plus d'une dimension alors x deviendra un vecteur représentant les coordonnées de l'individu.

$$x_i(t + 1) = x_i(t) + v_i(t + 1) \quad (1.1)$$

La force de l'algorithme réside donc dans l'expression de v et son efficacité à converger vers la solution optimale, en tenant compte des itérations précédentes et des relations entre individus. En effet la densité de la population est censée augmenter autour du point optimal.

C'est une famille d'algorithmes très vaste avec beaucoup de variantes où chaque type de problème peut avoir un algorithme associé. L'algorithme détaillé dans la section 1.3 fait partie de cette famille d'algorithmes.

1.1.2 Les algorithmes inspirés des colonies de fourmis ou Ant Colony Optimization (ACO)

L'optimisation par colonies de fourmis est l'un des algorithmes biologiques les plus connus, similaire à l'optimisation en essaim de particules et développé par Dorigo en 1996 [5]. L'idée générale vient de l'observation des colonies de fourmis à emprunter le chemin le plus court vers une source de nourriture. Elles étudient tous les chemins possibles, donc des solutions potentielles, et y déposent des phéromones dont la quantité est proportionnelle à la qualité du chemin suivi. Cette forme indirecte de communication est appelée "stigmergie" et intensifie les chemins d'intérêt. A cela s'ajoute que les phéromones peuvent s'évaporer, effaçant l'historique de recherche, pour permettre la diversification et trouver des chemins plus fructueux. Il y a alors une balance à évaluer entre les paramètres intensifs d'un côté et de diversité de l'autre pour permettre à l'algorithme de converger vers la solution optimale et de façon robuste. Ces deux aspects sont retrouvés dans la plupart des algorithmes biologiques [6, 7, 8]. Aucun paramètre

génétique n'est impliqué dans cette méthode, c'est à dire qu'aucune fourmi ne subira de mutation et/ou croisement, seule sa recherche compte.

L'application la plus courante concerne les problèmes combinatoires, souvent complexes à cause du nombre de solutions possibles [9]. Par exemple, le jeu de Pentamino, qui est un puzzle où le but est de contruire une figure à l'aide de pièces ayant différentes formes (figure 1.1).

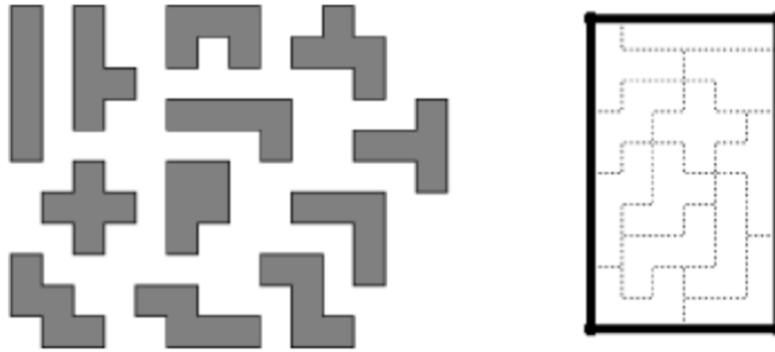


Figure 1.1: Figure représentant le jeu de Pentamino, à gauche les pièces du puzzle, à droite la figure à représenter [10]

Pour un nombre réduit de pentaminos, ces problèmes sont facilement résolubles. Cependant pour un nombre élevé, le nombre de possibilités devient difficilement calculable. L'algorithme permet alors de trouver le chemin le plus rapide entre un état initial et final, qui sera la construction de la figure désirée, en éliminant les chemins trop longs [10].

L'intérêt de résoudre ces problèmes va bien évidemment au-delà de ces puzzles. Dans l'industrie, les problématiques de planning, de production, de logistique ou de réseaux informatiques peuvent être optimisés via ces méthodes et ont démontré leur succès [11, 12, 13].

1.1.3 Les algorithmes basés sur les colonies d'abeilles ou Artificial Bee Colony Optimization (ABC)

Cet algorithme, développé par Karaboga en 2005, s'inspire de l'observation du comportement des abeilles à butiner [14]. Les abeilles sont divisées en trois classes :

- les employées : les abeilles qui ont une source de nourriture à exploiter
- les spectatrices : les abeilles qui jugent les danses des employées et choisissent les sources de nourriture en fonction de celles-ci
- les éclaireuses : les abeilles employées qui abandonnent leur source de nourriture pour en trouver une nouvelle

De même que précédemment, le nombre d'abeilles est égal au nombre de solutions de la population. Une seule abeille est assignée à une source et elle devient éclaireuse si elle l'abandonne.

Chaque source de nourriture est équivalente à une solution possible au problème d'optimisation et, par nature, ces individus ont tendance à être attirés par une source ayant une grande quantité de nourriture, c'est-à-dire, un meilleur résultat renvoyé par la fonction objectif. La différence ici est que si la quantité

de nectar à l'itération suivante est inférieure à celle précédemment trouvée, la position de l'employée ne changera pas. Une fois les nouvelles positions calculées, les informations sont transmises aux abeilles spectatrices. Elles vont pouvoir choisir les meilleures sources selon la densité en nourriture de chaque proposition. Cette décision repose sur un calcul de probabilité pondéré par la qualité des résultats, une bonne solution aura plus de chances d'être sélectionnée qu'une moins bonne. Enfin si une abeille ne parvient pas à améliorer sa position après un certain nombre d'itérations, elle abandonnera sa source pour devenir éclaireuse et chercher aléatoirement dans le domaine de définition.

Une variante de cet algorithme existe avec seulement trois abeilles, il a été développé par Rajasekhar en 2012 [15]. Dans cette version, la position de la meilleure abeille reste inchangée, la seconde explore les alentours de cette solution et la dernière est réinitialisée aléatoirement. Beaucoup d'autres variables ont été introduites pour le mouvement de recherche et montre l'évolution de la compréhension des méthodes d'optimisation car il implique moins d'individus pour un résultat équivalent.

1.1.4 L'évolution différentielle

L'idée fondamentale dans l'évolution différentielle est de créer de nouvelles solutions basées sur les solutions précédemment trouvées. Brièvement, les positions de trois individus x_1 , x_2 et x_3 choisis aléatoirement sont utilisées pour créer un nouvel individu "muté" (équation (1.2)).

$$v_i = x_1 + F(x_2 - x_3) \quad (1.2)$$

F est un facteur de proportionnalité qui donnera plus ou moins de "poids" vis-à-vis de la différence entre x_2 et x_3 . Ainsi un autre individu ayant une position x_i sera évalué de telle sorte que si sa valeur renvoyée $f(x_i)$, est moins intéressante que celle de l'individu "muté" $f(v_i)$, alors l'individu i sera remplacé par ce dernier. Dans le cas l'inverse, sa position x_i sera conservée. Cette méthode est appelée "croisement" ; la notion de génétique est alors introduite à travers ces algorithmes et sera utile par la suite [7].

On retrouve beaucoup d'applications à l'ingénierie chimique et pharmaceutique avec cette méthode, comme l'optimisation de l'émission de CO₂, la consommation en énergie des procédés ou encore la conversion de réactions complexes impliquant beaucoup d'équations [16, 17, 18].

1.1.5 Les algorithmes bactériens

Le premier algorithme bactérien ou Bacterial Foraging Optimization Algorithm (BFOA), a été proposé en 2002 par Passino [19]. Il s'inspire du comportement des bactéries, plus précisément *E. coli*, à rechercher de la nourriture et de leur propriétés chimiotactiques. Le comportement social d'une colonie de bactéries est intéressant à étudier d'un point de vue ingénierie, car il leur permet d'accéder aux meilleurs points de nourriture avec un minimum de risques. Une bactérie, ayant atteint une zone d'intérêt, communique sa position aux autres via des signaux chimiques, plus la quantité de nourriture est importante, plus le signal sera fort. De la même manière, pour une zone dangereuse (pauvre en nourriture ou présence d'un bactéricide), le signal indiquera de rester en dehors de celle-ci. L'algorithme mime alors quatre des principaux mécanismes d'une bactérie :

- chimiotactisme : Ce procédé simule le mouvement d'une bactérie via son flagelle. Biologiquement, elle peut se mouvoir de deux façons différentes : nager dans une direction pour une durée déterminée ou bien rouler sur elle-même. Elle alterne entre ces deux modes.
- essaimage : face à un point de nourriture, un groupe de bactéries va s'organiser en cercle et se déplacer en direction de celui-ci. Lorsque ces bactéries sont stimulées par du succinate, elles émettent de l'aspartate, une sorte d'appât qui les aide à s'agréger avec d'autres bactéries. Ainsi se créent des "essaims" concentriques, très denses en cellules, qui voyagent vers le point d'intérêt.
- reproduction : une bactérie peu saine peut mourir, tandis que les plus saines, mieux nourries, se reproduisent asexuellement par division cellulaire, donnant deux cellules à partir de la cellule mère. Cette caractéristique permet de garder la taille de l'essaim constant.
- élimination et dispersion : des changements graduels ou soudains de l'environnement peuvent compromettre la survie des bactéries, comme une hausse de température, même si cet environnement contient beaucoup de ressources. Ce type d'évènement peut tuer certains individus ou bien les disperser. Pour simuler de tels phénomènes, certaines de ces cellules sont tuées aléatoirement avec une faible probabilité et aléatoirement remplacées sur le domaine de recherche.

La partie génétique est ici très présente et permet une exploration large du domaine de définition. Les applications industrielles sont similaires à celles retrouvées pour les algorithmes d'évolution différentielle [20].

1.2 Comparaison avec les méthodes classiques

D'autres algorithmes ont vu le jour plus récemment et reprennent certaines caractéristiques des méthodes présentées, comme par exemple, le clignotement des lucioles, le déplacement des grenouilles, ou encore l'écholocation des chauves-souris [21, 9, 22]. Un exemple sera présenté en détail dans la partie 1.3. Ce sont des méthodes de résolution particulièrement performantes pour des systèmes composés de beaucoup de dimensions. Grâce à la puissance de calcul d'un ordinateur commun, il est alors possible de paralléliser facilement les tâches, c'est à dire effectuer un calcul simultané pour un ensemble de points, ici une population d'individus. On notera tout de même que tous ces algorithmes biologiques présentés impliquent une grande part d'aléatoire ; ainsi il est toujours judicieux de lancer plusieurs fois le programme pour un même problème afin de diminuer les risques de convergence précoce [23].

Par comparaison avec les algorithmes traditionnels basés sur le gradient, comme la méthode de quasi-Newton, ceux-ci demandent beaucoup plus de ressources et reposent sur des bases mathématiques relativement complexes [24]. Leur principal inconvénient vient de l'initialisation, ou du point de départ, du calcul car leur approche est locale et non globale. En effet, même si l'erreur au point optimal est comparable voire meilleure que celle des algorithmes biologiques, ils risquent de rester enfermés dans des extrema locaux de la fonction considérée [25]. Ainsi grâce à la "globalité" des bio-algorithmes, la probabilité de passer à côté de l'optimum recherché est plus faible. Cet aspect sera traité plus en détail dans le paragraphe 1.3.8.

Pour finir, il est possible de combiner ces deux approches de sorte qu'un algorithme global initialise un algorithme local. On s'affranchit alors de la plupart des inconvénients propres à chacun d'eux, avec un bon point de départ et une convergence forte pour obtenir des résultats comparables voire supérieurs

[26]. Ils sont appelés "algorithmes hybrides" et sont reconnus pour la qualité des résultats sur des fonctions usuelles de performance. En pratique, une telle précision n'est pas utile et n'a de sens que pour la recherche [27, 28].

1.3 L'algorithme Krill Herd

L'algorithme utilisé ici est un algorithme en essaim appelé Krill Herd Algorithm (KHA), combinant les avantages de l'intelligence en essaim et des algorithmes génétiques [29]. Cette méthode est basée sur la simulation de bancs de krills en réponse à un processus environnemental. Presque tous les paramètres de cet algorithme sont obtenus à partir d'études de cas réels. La fonction d'"adaptabilité" de chaque krill, c'est à dire la façon dont il va se mouvoir, est définie selon sa distance d'un point de nourriture et sa distance par rapport à l'essaim qui possède la densité la plus forte. Ainsi la position de chaque organisme est régulée selon :

1. Le mouvement induit par les autres krills
2. Sa recherche de nourriture
3. Sa diffusion aléatoire

Le regroupement des espèces marines n'est pas un phénomène totalement aléatoire et/ou diffus, les principaux mécanismes identifiés sont liés à leur capacité à se nourrir, se reproduire, se protéger des prédateurs et s'adapter aux conditions environnementales. Le krill d'Antarctique est l'une des espèces marine les plus étudiées en particulier pour sa caractéristique à former de larges essaims. Il reste toutefois des zones d'ombre sur les forces déterminant la distribution des krills. Lorsque les prédateurs (phoques, mouettes ou bien pingouins) attaquent ces organismes, ils suppriment un krill, diminuant la densité globale de l'essaim. Après l'attaque, la reformation de l'essaim va dépendre de deux objectifs :

1. Augmenter sa densité en krills
2. Accéder à de la nourriture

Ces objectifs vont conduire la population à se rassembler au plus proche d'un minimum global. Plus la distance d'un krill vis-à-vis du meilleur point de nourriture sera courte et la densité de son essaim grande, plus la fonction objectif approchera du point optimal.

1.3.1 Modèle Lagrangien

La prédation a été présentée comme responsable de la baisse de la densité de l'essaim et de l'éloignement du point de nourriture. Ce procédé est considéré comme la phase d'initialisation de l'algorithme. L'adaptabilité est, elle, quantifiée selon la valeur de la fonction objectif. Afin de pouvoir traiter des problèmes d'optimisation ayant n dimension(s), le modèle Lagrangien suivant est généralisé par l'équation (1.3) où N_i est la quantité de mouvement induit par les autres krills, F_i celle de la recherche de nourriture et D_i la diffusion physique du i^{me} krill.

$$\frac{dX_i}{dt} = N_i + F_i + D_i \quad (1.3)$$

Ce modèle définit alors, comme présenté dans la partie 1.1.1, la direction et la quantité de mouvement appliqué à chaque individu et à chaque itération sous l'influence de ces trois composantes.

1.3.2 Mouvement induit par les autres krills

La direction du mouvement induit, α_i , est estimée à partir de l'effet de densité local, de l'effet de densité cible et la répulsion des krills entre eux. Ceci peut être défini par l'équation (1.4) et (1.5).

$$N_i^{new} = N^{max} \alpha_i + \omega_n N_i^{old} \quad (1.4)$$

$$\alpha_i = \alpha_i^{local} + \alpha_i^{target} \quad (1.5)$$

N^{max} étant la vitesse maximale d'un krill, N_i^{old} est le mouvement induit de l'itération précédente, ω_n le poids apporté à cette dernière 0 et 1 aléatoirement, α_i^{local} est l'effet local des voisins et α_i^{target} est la direction qui cible le krill ayant la meilleure position. La vitesse maximale N^{max} est une constante mesurée et égale à 0.01 m/s. L'effet des voisins α_i^{local} sur un individu i peut être attractif ou répulsif, il est déterminé selon les équations (1.6), (1.7) et (1.8).

$$\alpha_i^{local} = \sum_{j=1}^{NN} \hat{K}_{ij} \hat{X}_{ij} \quad (1.6)$$

$$\hat{X}_{ij} = \frac{X_j - X_i}{\|X_j - X_i\| + \varepsilon} \quad (1.7)$$

$$\hat{K}_{ij} = \frac{K_i - K_j}{K^{worst} - K^{best}} \quad (1.8)$$

K^{best} et K^{worst} sont respectivement la meilleure et la pire des valeurs de la fonction objectif parmi tous les individus, K_i la valeur de la fonction pour le i^{eme} krill, K_j celle de son voisin, X représente la positions et NN le nombre total de voisins. Pour éviter certaines singularités, un petit nombre positif ε est ajouté au dénominateur.

Pour choisir un voisin, plusieurs stratégies peuvent être employées. Par exemple, l'introduction d'un "ratio de voisinage" qui permet de compter les individus les plus proches d'un certain krill. Dans ce cas, une distance de détection d_s doit être appliquée autour de tous les individus pour définir l'identité et le nombre de voisins de chacun. Cette dernière est modélisée selon l'équation (1.9).

$$d_{s,i}^{local} = \frac{1}{5N} \sum_{j=1}^N \|X_i - X_j\| \quad (1.9)$$

Le facteur 5 du dénominateur est obtenu empiriquement, N représente le nombre total de la population et $d_{s,i}$ la distance de détection appliquée au i^{eme} krill. Ainsi, si la distance entre deux individus i et j est inférieure à cette distance calculée, ils sont considérés comme voisins (figure 1.2).

La cible α_i^{target} de chaque krill est la plus basse valeur de la fonction obtenue par l'un des krills dans l'entière population. L'effet de cette position sur le krill est formalisé selon l'équation (1.10).

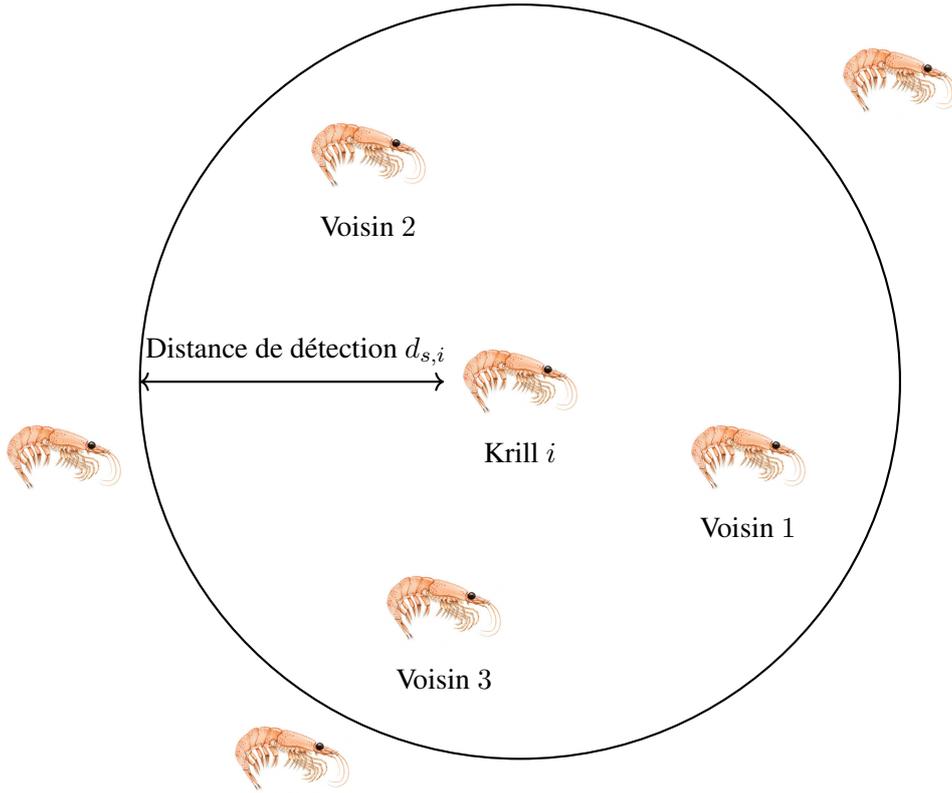


Figure 1.2: Schéma représentatif du principe de la distance de détection appliqué à chaque krill pour déterminer son voisinage [29]

$$\alpha_i^{target} = C^{best} \hat{K}_{i,best} \hat{X}_{i,best} \quad (1.10)$$

C^{best} est un coefficient (ou un poids) qui possède une partie aléatoire, $rand$ défini entre 0 et 1, permettant l'exploration et une partie variable I , actuel numéro d'itération ainsi que I_{max} le nombre maximum d'itérations définies (équation (1.11)).

$$C^{best} = 2 \left(rand + \frac{I}{I_{max}} \right) \quad (1.11)$$

Ce coefficient ne peut donc être supérieur à 4 et tend à évoluer positivement au fur et à mesure des itérations. Il permet d'intensifier la valeur cible renvoyée par le meilleur individu lorsque l'algorithme approche de la dernière itération, l'influence de α_i^{target} devient alors de plus en plus forte sur le mouvement de la population.

1.3.3 Recherche de nourriture

Le mouvement induit par la nourriture peut être décomposé en deux paramètres, le premier étant la localisation du point de nourriture et le second sa valeur à l'itération précédente (équation (1.12) et (1.13)).

$$F_i = V_f \beta_i + \omega_f F_i^{old} \quad (1.12)$$

$$\beta_i = \beta_i^{food} + \beta_i^{best} \quad (1.13)$$

V_f étant la vitesse de recherche, F_i^{old} la dernière valeur de ce mouvement, ω_f le poids appliqué à ce dernier (compris entre 0 et 1 aléatoirement), β_i^{food} l'attraction vers le meilleur point de nourriture et β_i^{best} l'effet de la meilleure valeur obtenue par le i^{eme} krill. La vitesse de recherche V_f est définie comme étant égale à 0.02 m/s. L'attraction est variable selon la localisation du meilleur point de nourriture trouvé parmi toute la population, car il est bien évident que la valeur optimale ne peut être fournie en premier lieu. Comme il n'est pas possible de la déterminer directement, la valeur optimale est estimée grâce à la distribution de tous les krills sur la fonction objectif, comme un centre de masse (équation (1.14))

$$X^{food} = \frac{\sum_{i=1}^N \frac{1}{K_i} X_i}{\sum_{i=1}^N \frac{1}{K_i}} \quad (1.14)$$

Au final, l'attraction ressentie par le i^{eme} krill soumis à ce centre de masse peut être déterminée selon les équations (1.15) et (1.16) et conduit les krills vers l'optimum global après quelques itérations.

$$\beta_i^{food} = C^{food} \hat{K}_{i,food} \hat{X}_{i,food} \quad (1.15)$$

$$C^{food} = 2 \left(1 - \frac{I}{I_{max}} \right) \quad (1.16)$$

A l'inverse de la relation (1.11) pour l'effet de densité, C^{food} est un coefficient qui diminue l'effet de la nourriture au fur et à mesure des itérations jusqu'à même l'annuler. L'intérêt est de laisser place à l'exploration en fin de processus au cas où le centre de masse dirigerait tous les individus vers un piège et préserve ainsi la "globalité" de l'algorithme. Enfin, l'effet de la meilleure position, donnée par $K_{i,best}$, de chaque krill a également son importance sur leur propre mouvement (équation (1.17)).

$$\beta_i^{best} = \hat{K}_{i,best} \hat{X}_{i,best} \quad (1.17)$$

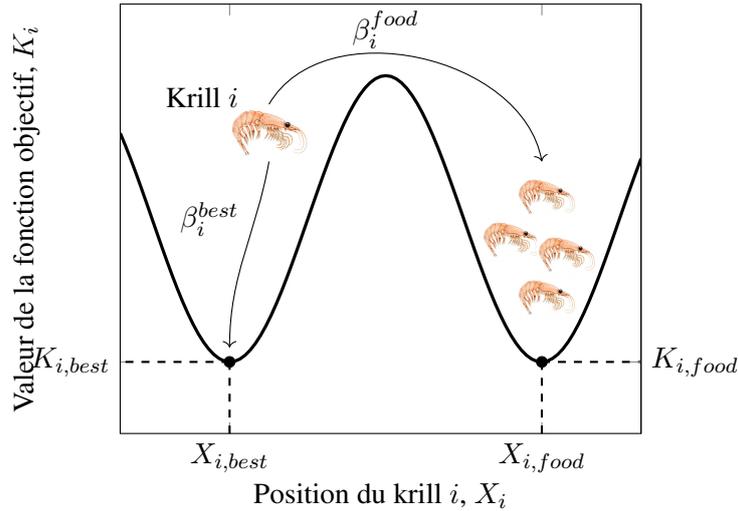


Figure 1.3: Schéma représentatif de l'effet de β_i , d'un côté la composante β_i^{best} qui attire le krill i vers sa meilleure position visitée, et β_i^{food} l'effet de centre de masse

La figure 1.3 montre la balance entre ces deux effets : les individus évitent de tous converger vers le centre de masse (le point de nourriture noté ici "food") ou de stationner sur leur meilleure valeur grâce à un dosage évolutif entre ces deux composantes. Au final, le paramètre β est l'une des caractéristiques les plus influentes de cet algorithme avec α ; à eux deux ils participent en grande partie à sa capacité de convergence vers l'optimum.

1.3.4 Diffusion physique

Le phénomène diffusif des organismes est considéré comme principalement aléatoire, il est composé d'une vitesse maximale aléatoire D^{max} comprise entre 0.002 et 0.01 m/s et de δ , vecteur aléatoire dont les valeurs sont comprises entre -1 et 1 (équation (1.18)).

$$D_i = D^{max} \delta \quad (1.18)$$

Seulement, il est préférable d'adapter cette relation de telle sorte que, plus un krill sera en position favorable, c'est à dire proche de l'optimum, moins ce paramètre aura d'impact. Donc un nouveau terme est ajouté à l'équation (1.18) pour prendre en compte cet effet et, de la même manière que précédemment avec la relation (1.16), plus on avance dans les itérations, moins l'effet de diffusion sera fort (équation (1.19)).

$$D_i = D^{max} \left(1 - \frac{I}{I_{max}} \right) \delta \quad (1.19)$$

Aucune direction préférentielle n'est donc donnée par ce terme diffusif, rendant le comportement des individus moins prévisible, du moins pour les premières itérations.

1.3.5 Modèle itératif de l'algorithme

Les mouvements définis précédemment tentent de conduire chaque organisme vers l'optimum global, grâce à l'effet de densité et la recherche de nourriture qui contiennent chacun un effet global et un effet local. Ils fonctionnent en parallèle et contribuent à l'efficacité de l'algorithme. En d'autres termes, si

l'une des valeurs K_j , K^{best} , K^{food} et K_i^{best} , est meilleure que celle renvoyée par la position du krill i , alors il y aura un effet attractif. Dans le cas contraire, il sera répulsif. Le vecteur comportant les nouvelles coordonnées d'un krill i en tenant compte de tous les effets présentés précédemment et pendant l'intervalle de temps Δt est donné par l'équation (1.20).

$$X_i(t + \Delta t) = X_i(t) + \Delta t \frac{dX_i}{dt} \quad (1.20)$$

Δt est une constante importante et doit être adaptée au problème à optimiser, car elle est directement liée aux vecteurs vitesse des individus. Elle dépend de l'espace de recherche considéré et peut être obtenue suivant l'équation (1.21).

$$\Delta t = C_t \sum_{j=1}^{NV} (UB_j - LB_j) \quad (1.21)$$

NV étant le nombre total de variables, LB_j et UB_j sont respectivement les bornes basses et hautes de chaque variable j . La soustraction de ces bornes donne l'espace de définition et C_t est une constante empirique comprise entre 0 et 2. Une valeur basse de C_t permettra au krill d'explorer l'espace avec prudence.

1.3.6 Les opérateurs génétiques

Pour améliorer la performance de l'algorithme, des mécanismes de reproduction génétique sont incorporés juste après la mise à jour des nouvelles positions décrites par la relation (1.20). Il y en a deux : la mutation et le croisement.

Le croisement

D'abord employé dans les algorithmes évolutionnistes, l'opérateur de croisement est une stratégie efficace pour le travail d'optimisation. Il est contrôlé par une probabilité de croisement Cr . En générant aléatoirement un nombre entre 0 et 1, le m^{eme} composant de X_i , $x_{i,m}$, est manipulé selon le système conditionnel présenté en (1.22) et (1.23).

$$x_{i,m} = \begin{cases} x_{r,m} & rand_{i,m} < Cr ; \quad r \in \{1, 2, \dots, i-1, i+1, \dots, N\} \\ x_{i,m} & else \end{cases} \quad (1.22)$$

$$Cr = 0.2 \tilde{K}_{i,best} \quad (1.23)$$

Ainsi, tous les individus de la population ont une probabilité de voir au moins une de leurs coordonnées remplacée par celle d'un autre individu r choisi au hasard. Ceux au plus proche de l'optimum global seront épargnés de tout changement.

La mutation

De la même manière, le phénomène de mutation est contrôlé par une probabilité propre, Mu . La mutation adaptative est formulée par le système conditionnel décrit en (1.24) et (1.25).

$$x_{i,m} = \begin{cases} x_{gbes,m} \cdot \mu \cdot (x_{p,m} - x_{q,m}) & rand_{i,m} < Mu ; p, q \in \{1, 2, \dots, i-1, i+1, \dots, K\} \\ x_{i,m} & else \end{cases} \quad (1.24)$$

$$Mu = 0.05 \hat{K}_{i,best} \quad (1.25)$$

Contrairement au croisement, la mutation n'échange pas les valeurs entre individus mais les change pour une nouvelle valeur aléatoire, notamment grâce à μ qui est compris entre 0 et 1. La probabilité de mutation auprès de l'optimum global est donc faible et augmente lorsque l'on s'éloigne de ce point.

Intérêt des opérateurs génétiques

Ces opérateurs sont particulièrement intéressants lorsque la fonction objectif présente de multiples minima locaux de coordonnées proches. On peut comparer ces minima à des pièges pour l'algorithme car il y a un risque de converger trop rapidement vers l'un d'eux selon le calcul du mouvement (équation (1.3)). Ainsi, les krills les plus éloignés de la meilleure solution trouvée ont une forte probabilité de subir ce brassage aléatoire. Cela permet l'exploration de nouvelles zones de la fonction tout en gardant les meilleures positions occupées par une partie de la population. Bien évidemment, chaque brassage ne garantit pas de trouver un meilleur résultat, mais donne une chance à l'algorithme de sortir de sa position si une solution optimale existe ailleurs.

A titre d'exemple, la figure 1.4 présente graphiquement la fonction de Rastrigin : il s'agit d'une fonction ayant de multiples minima locaux très rapprochés les uns des autres et espacés de maxima locaux. La fonction sera également présentée dans la section 1.3.8 via l'équation (1.27). Ce type de fonction se destine à évaluer les performances des algorithmes d'optimisation (fonction dite de "benchmark").

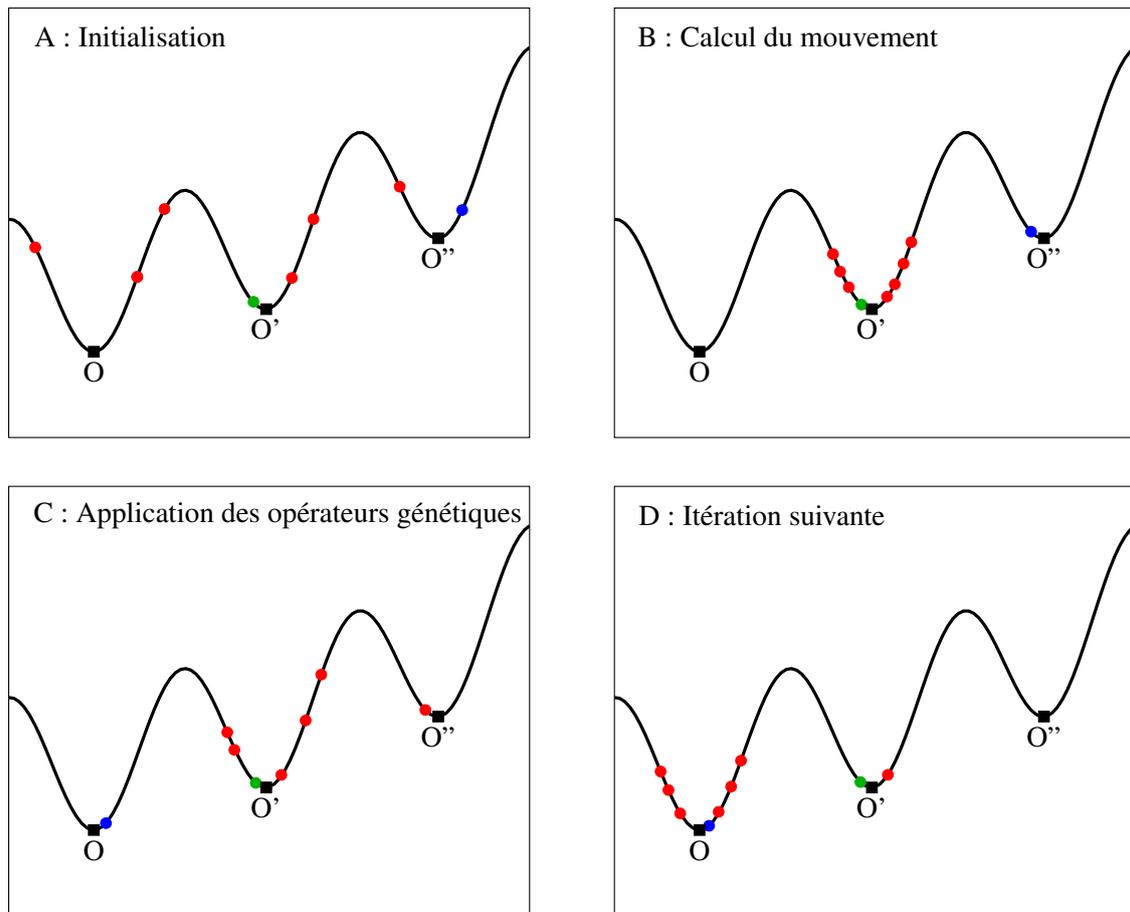


Figure 1.4: Schéma illustrant un exemple de situation d'intérêt des opérateurs génétiques, en bleu l'individu muté ou croisé, en vert l'individu ayant la meilleure valeur avant mutation ou croisement

Pour cet exemple, on utilisera cette fonction avec une seule dimension, c'est à dire pour $n = 1$. Après une initialisation aléatoire (graphique A), la meilleure solution renvoyée (en vert) conduit les krills vers un minimal local O' . Le calcul du mouvement concentre la population autour de ce point (graphique B) et rend difficile l'exploration en-dehors de cette zone. Un krill, relativement éloigné de cette solution, est alors muté ou croisé (en bleu) de façon aléatoire près du point O devenant ainsi l'individu présentant la meilleure solution de l'itération (graphique C). Cette position va naturellement induire un déplacement de la population à l'issue de l'itération suivante (graphique D) et permettre la convergence vers l'optimum global qui est le point O . Le cas ainsi présenté montre l'intérêt de tels opérateurs ; cependant il faut noter que de telles situations sont rares et peuvent parfois être néfastes sur la qualité du résultat final à cause de mutations ou croisements trop intempestifs. C'est pourquoi les critères Cr et Mu doivent être soigneusement adaptés selon la complexité du problème considéré.

Pour finir, il est aussi possible dans certains cas simples, de s'affranchir de ces mimétismes génétiques. L'initialisation aléatoire d'une population conséquente sur un domaine de définition relativement restreint peut être suffisante pour approcher directement l'optimum global de la fonction objectif, diminuant ainsi le temps de calcul total.

1.3.7 Critère d'arrêt, méthodologie et programme

La question d'un critère d'arrêt se pose lors de la composition de cet algorithme. Dans un premier temps, il est obligatoire de fixer un nombre maximal d'itérations car les trois facteurs de l'algorithme, N_i , F_i et D_i , en dépendent. Afin de réduire le temps de calcul, il pourrait être judicieux de s'arrêter avant cette limite. En effet, un critère d'arrêt simple consisterait à comparer la valeur minimisée de l'itération précédente par rapport à la nouvelle valeur calculée et d'arrêter le processus si ces deux valeurs sont très proches. Cependant, l'instauration d'un critère d'arrêt pourrait perturber le bon fonctionnement de l'algorithme si l'on se réfère aux relations (1.11), (1.16) et (1.19). Un krill peut rester dans un minimum local pendant plusieurs itérations puis se déplacer par la suite si une meilleure solution est découverte à l'approche de la dernière itération. L'instauration d'un tel critère augmenterait alors la probabilité de ne pas renvoyer la valeur optimale. C'est pourquoi il est préférable d'effectuer toutes les itérations prévues et de sélectionner la meilleure valeur parmi tous les résultats de tous les individus et pour toutes les itérations.

Un logigramme simplifié de cet algorithme est présenté en annexe 1 et illustre la méthodologie employée pour l'élaboration du programme d'optimisation. Ce programme a donc été spécialement développé et inclut cet algorithme sur base de ce qui a été présenté dans la partie 1.3. L'ensemble du code, en Fortran 90, est présenté en annexe 2 de ce document.

1.3.8 Fonctions d'évaluation de la performance de l'algorithme

Ce programme permet dans un premier temps de proposer à l'utilisateur une série de fonctions à minimiser ainsi que d'effectuer un calcul d'équilibres de phases qui sera présenté dans le chapitre 2. Pour tester les performances de l'algorithme, deux fonctions d'évaluation sont présentées :

- la fonction d'Ackley (équation (1.26))
- la fonction de Rastrigin (équation (1.27))

$$A(X) = -20 \exp \left(-0.2 \sqrt{0.5 \sum_{i=1}^n x_i^2} \right) - \exp \left(0.5 \sum_{i=1}^n \cos(2\pi x_i) \right) + \exp(1) + 20 \quad (1.26)$$

$x_i \in [-32, 32], i \in [1, 2, \dots, n-1, n]$

$$R(X) = 10n + \sum_{i=1}^n (x_i^2 - 5 \cos(2\pi x_i)) \quad (1.27)$$

$x_i \in [-5.12, 5.12], i \in [1, 2, \dots, n-1, n]$

De multiples minima locaux sont retrouvés sur leurs domaines de définition, comme présenté sur la figure 1.5 afin de tester les performances de l'algorithme.

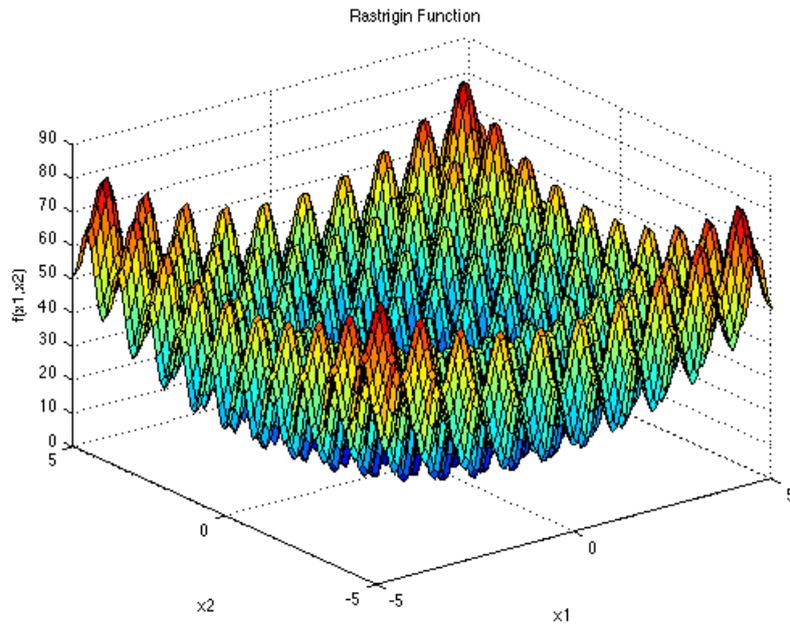
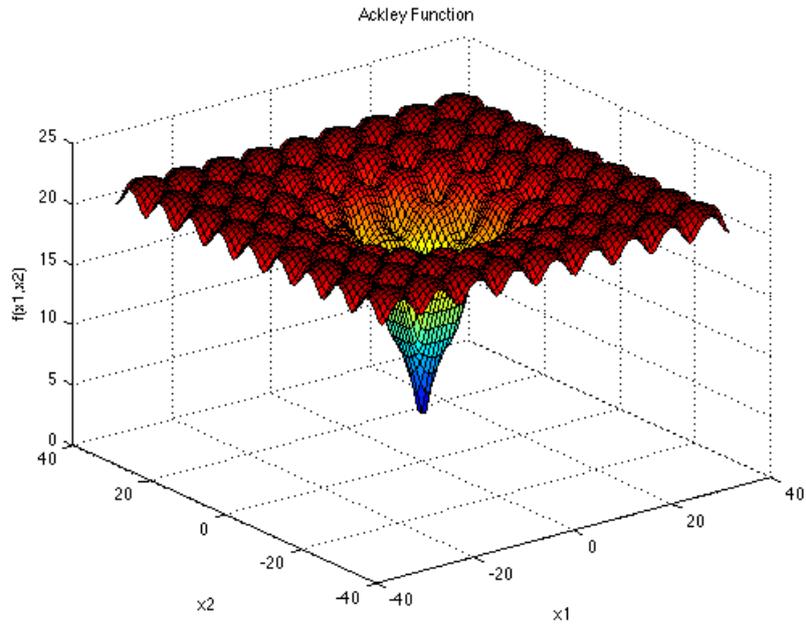


Figure 1.5: Graphiques des fonctions d'Ackley (en haut) et de Rastrigin (en bas) pour $n = 2$, le minimum global des deux fonctions est 0 [30, 31]

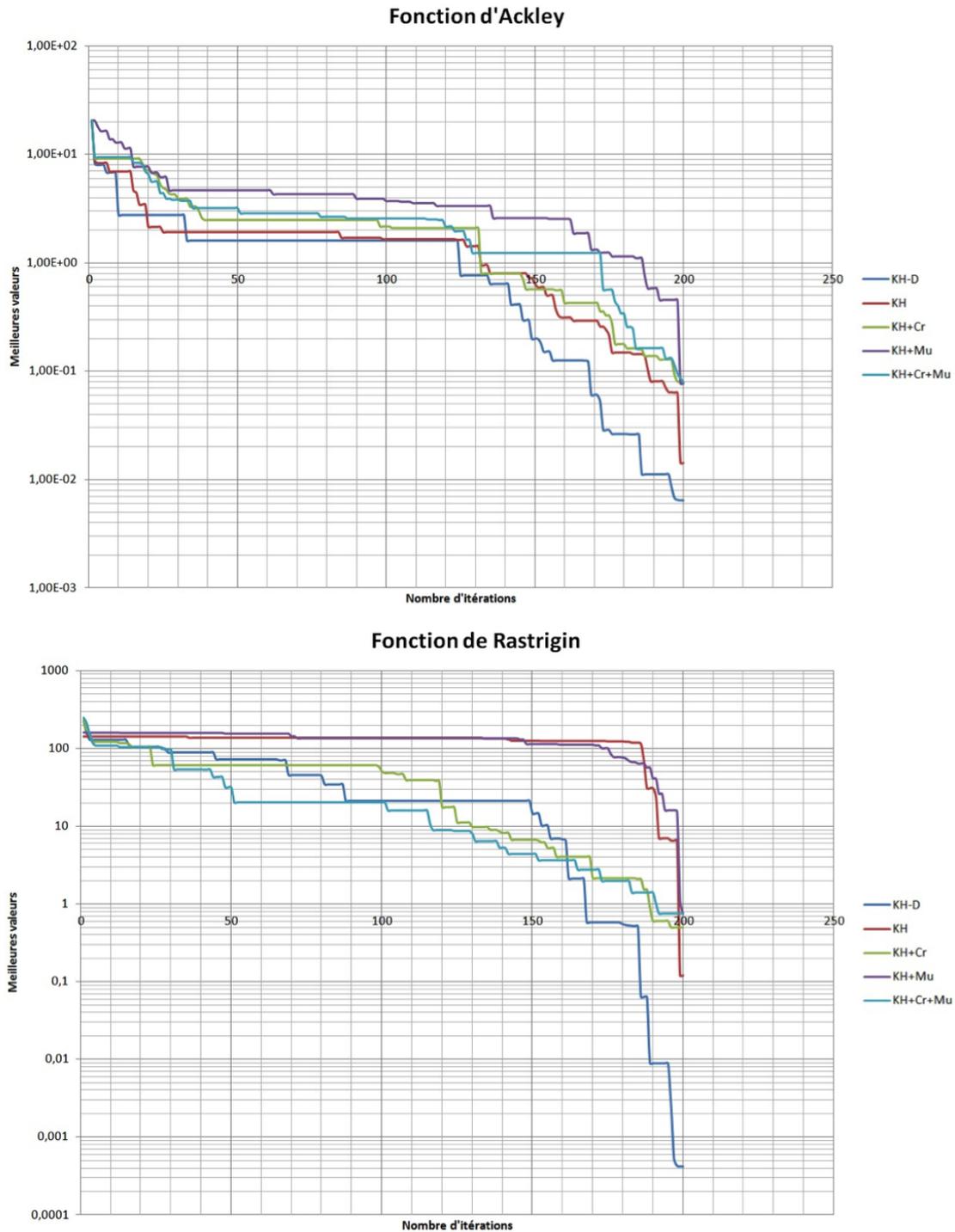


Figure 1.6: Graphique semi-logarithmique des valeurs renvoyées par l’algorithme Krill Herd au fur et à mesure des itérations pour la fonction d’Ackley (en haut) et de Rastrigin (en bas), le nombre de dimensions est fixé à $n = 20$

En plus de l’algorithme complet, quatre variantes sont étudiées pour chacune des fonctions (figure 1.6), sans opérateurs génétiques ni diffusion (en bleu), sans opérateurs génétiques et avec diffusion (en rouge), avec croisement et diffusion (en vert), avec mutation et diffusion (en violet) et enfin l’algorithme complet (en bleu clair). Premièrement, comme évoqué dans la section 1.3.6, on peut voir clairement que

le meilleur résultat final est obtenu lorsque toutes les composantes aléatoires sont retirées. Deuxièmement, la mutation induit une convergence lente vers l'optimum global, elle interfère négativement avec la recherche sur le domaine et présente des résultats finaux moins satisfaisants, ce paramètre pourrait alors être mieux ajusté, voire retiré. Troisièmement, même si le croisement est similaire à la mutation, la convergence est plus rapide et limite l'effet de piège. Dans tous les cas, l'optimum global est approché à chaque fois et pour de telles fonctions, les résultats sont de l'ordre de 10^{-1} à 10^{-3} , ce qui est amplement satisfaisant sachant que les applications réelles sont bien moins complexes.

Le chapitre suivant développera plus particulièrement la fonction objectif choisie pour l'utilisation de l'algorithme. Le but final étant de construire un diagramme de phases pour un système binaire de solvants impliqués dans une opération de distillation industrielle.

2 Construction d'un diagramme de phases à l'aide d'un bioalgorithme

Gilbert Newton Lewis était un physicien et chimiste célèbre qui avait pour habitude de dire : "Thermodynamics exhibits no curiosity". En d'autres termes, les conclusions tirées de la thermodynamique restent plutôt générales. En effet, si on applique les règles de la thermodynamique à un système en équilibre, le résultat sera vrai même si la nature des molécules impliquées n'est pas connue. Pour cette raison, il n'y a que très peu d'équations spécifiques et celles-ci peuvent être employées pour de nombreuses applications. Par exemple, la formalisation des équations impliquées dans les équilibres liquide-vapeur et les réactions chimiques sont au final similaires [32].

Le problème considéré ici concerne un calcul de flash, ou vaporisation d'un liquide, à température et pression constantes par minimisation de l'énergie de Gibbs de mélange. Dans un premier temps, des notions générales de thermodynamique seront introduites pour caractériser et comprendre les paramètres en jeu. Dans un second temps, pour résoudre numériquement les équations impliquées dans les équilibres liquide-vapeur, des modèles empiriques et mécanistiques seront présentés et utilisés pour pouvoir finalement construire un diagramme de phase. Il existe une grande variété de modèles mécanistiques plus ou moins adaptés selon la polarité des composés, la pression ou encore la température du système. Le but ici n'est pas de tous les étudier, mais de choisir des modèles pertinents qui, au final, aboutiront à une bonne représentation du mélange considéré.

2.1 Calcul de flash par minimisation de l'énergie de Gibbs de mélange

2.1.1 Première et deuxième loi de la thermodynamique

La thermodynamique est une branche de la physique qui traite de l'énergie d'un système. Même s'il est facile de se représenter ce qu'est l'énergie, il est difficile de lui donner une définition précise. L'énergie peut être perçue aussi bien comme un travail ou comme une aptitude à modifier des systèmes. D'après la première loi fondamentale, l'énergie d'un système est conservée dans tous les procédés thermodynamiques. C'est à dire qu'elle est capable de changer d'un état vers un autre mais le total reste inchangé [33, 34, 35].

Cependant, un procédé qui respecte cette première loi n'est pas une condition suffisante pour l'observer dans la nature. Par exemple, si un objet chaud est mis en contact avec un objet froid, le transfert d'énergie sous forme de chaleur s'effectuera du point chaud vers le point froid jusqu'à atteindre un équilibre. Le scénario inverse où le transfert de chaleur irait de l'objet froid vers l'objet chaud est impossible mais l'énergie serait tout de même conservée, la première loi est respectée. D'où l'importance d'introduire une variable d'état S , ou entropie, pour formuler le second principe. L'entropie peut être considérée comme le désordre statistique des molécules d'un système. Ainsi, une variation d'entropie δS est égale au transfert de chaleur δQ sur la température T comme présenté par l'équation (2.1) [36].

$$\delta S = \frac{\delta Q}{T} \quad (2.1)$$

Entre un état initial i et final f , l'entropie totale d'un système et celle de son environnement est inchangée si le phénomène est réversible, soit $S_i = S_f$. Dans le cas des deux objets en contact, le phénomène est irréversible [37].

2.1.2 Les conditions d'équilibre entre phases

Il n'est pas rare que pour certaines conditions de T et P , un système fermé renfermant un corps pur soit sous forme de plusieurs phases en équilibre. Ainsi, le système polyphasique est en équilibre si et seulement si :

- Il n'y a plus de transfert de chaleur entre les phases : la température T est uniforme dans le système
- Il n'y a plus d'échange de travail volumique entre les phases : la pression P est uniforme dans le système
- Il n'y a plus d'échange de matière entre les phases : le potentiel chimique μ de chaque constituant est uniforme dans le système

La notion de potentiel chimique est intimement liée à l'énergie de Gibbs G , sa dimension est une énergie molaire. Le potentiel chimique est défini comme la différentielle de l'énergie de Gibbs à T et P fixées. Il représente donc l'énergie libre de Gibbs molaire g , si l'on considère n comme la quantité molaire du composé considéré (equation (2.2)).

$$\mu(T, P) = \left(\frac{\partial G}{\partial n} \right)_{(T, P)} = g(T, P) \quad (2.2)$$

En réalité beaucoup de définitions existent et relient le potentiel chimique avec d'autres fonctions d'état. C'est bien l'énergie libre de Gibbs qui doit être choisie ici si l'on veut étudier un système à pression et température constantes (isobare et isotherme). Ce sont les conditions les plus retrouvées dans l'industrie pharmaceutique pour des opérations de distillation. A titre d'exemple, pour un système isochore (volume constant) et isotherme, c'est l'énergie libre de Helmholtz F qui serait choisie en tant que fonction d'état. La définition de potentiel chimique permet donc d'introduire la notion d'énergie de Gibbs qui sera essentielle pour la suite [34, 38, 33].

2.1.3 Energie libre de Gibbs

Dans cette partie, on s'intéressera aux propriétés de l'énergie libre de Gibbs, ou enthalpie libre, d'un système qui gouvernent sa stabilité thermodynamique, tout particulièrement des conditions de séparation de phases.

L'énergie libre de Gibbs est une fonction d'état qui implique l'enthalpie H , l'entropie S et la température T . Elle s'exprime en joules et est définie par l'équation (2.3).

$$G = H - TS \quad (2.3)$$

Cette énergie est particulièrement utile pour définir les conditions d'équilibre nécessaires aux réactions à température et pression constantes [33]. Un procédé spontané libère l'énergie nécessaire pour se réaliser, diminuant l'énergie libre de Gibbs de l'état initial à final. En d'autres termes, la différence d'énergie

libre de Gibbs pour une réaction spontanée doit être négative [39].

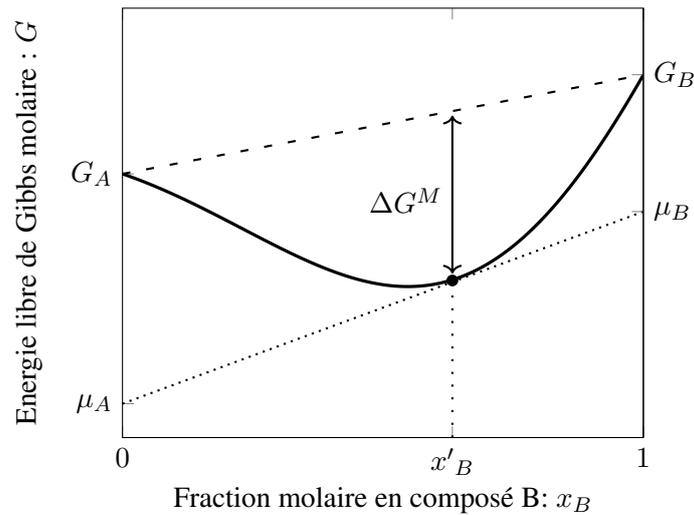


Figure 2.1: Energie libre de Gibbs molaire G d'un système binaire stable en fonction de la fraction molaire en composé B

On considère par exemple un mélange de deux composés A et B , G_A représente l'énergie de Gibbs du composé A pur et G_B celle du composé B pur. A température et pression constantes, l'énergie de Gibbs molaire d'un système doit être minimisée lorsque les deux composés sont mis en contact. Sur la figure 2.1, la courbe pleine représente la variation de l'énergie de Gibbs molaire du mélange en fonction de la composition en composé B . La droite en pointillés entre G_A et G_B représente cette même variation mais pour un système non mélangé ou idéal. La distance entre ces deux courbes représente la variation d'énergie due au mélange. Ici, l'énergie libre de Gibbs varie de façon monotone avec la composition, dans le sens où en chaque point, l'énergie G de la solution est toujours inférieure à la droite $[G_A ; G_B]$. Par conséquent, il n'y aura qu'une seule phase stable : la solution. Si l'inverse était vrai, il y aurait une séparation en deux phases liquides [35, 39].

On considère maintenant un autre exemple où l'intervalle de composition considéré x permet aussi une séparation en deux phases distinctes (figure 2.2). L'ajout de B dans A pur (et inversement) induit d'abord une décroissance de G jusqu'à atteindre deux minima locaux, les points O et O' . Cette fois-ci, l'énergie présente un maximum local M entre ces deux points, lorsque les fractions molaires de chaque composé sont à peu près équivalente dans cet exemple.

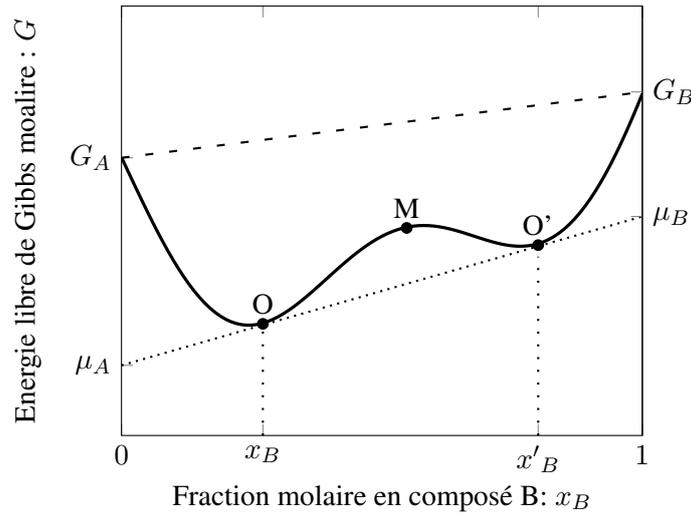


Figure 2.2: Variation de l'énergie libre de Gibbs molaire G d'un système binaire en fonction de la fraction molaire en composé B

Il faut noter ici que chaque valeur réelle de G est bien inférieure à la droite $[G_A ; G_B]$ et donc comme précédemment, il n'y aurait donc pas de séparation de phases. Cependant, la droite $[O ; O']$ représente l'énergie libre de Gibbs molaire d'un système séparé en deux phases, l'une de composition x_B l'autre de composition x'_B . Ce système partiellement démixé a une énergie plus faible que la courbe G et, par conséquent, une composition au point M va spontanément se séparer en deux phases coexistantes ayant pour compositions respectivement x_B et x'_B . Pour rappel, il est essentiel que le potentiel chimique des deux phases soit identique, c'est un critère d'équilibre [39]. A l'inverse, pour des compositions inférieures à x_B et supérieures à x'_B , une seule phase sera observée.

Il est possible d'aller plus loin dans l'analyse graphique de la figure 2.2, notamment la définition des domaines métastables, mais l'objectif ici est de définir uniquement les conditions qui impliquent la séparation ou non des phases d'un système stable. Ces conditions sont nécessaires à la construction d'un diagramme de phases. Lorsque température et pression sont les variables indépendantes d'un système à un composé, les phases stables (solides, liquides ou vapeurs) sont celles qui possèdent l'énergie libre de Gibbs totale la plus basse. Ce principe est aussi applicable à un système binaire et au delà [38].

2.1.4 Grandeurs de mélange et solutions réelles liquides

Pour caractériser les solutions réelles, comme aucune solution n'est idéale, il faut s'intéresser aux grandeurs d'excès. Une grandeur d'excès, ou de mélange, est la différence entre la grandeur réelle totale, avec celle de la solution idéale correspondante pour une température, une pression et une composition données. Elle quantifie alors la déviation de la solution par rapport à l'idéalité de façon positive ou négative (voir figure 2.1 et 2.2). Ainsi, on peut relier G l'énergie libre de Gibbs totale, G^M l'énergie libre de Gibbs de mélange et G^{id} l'énergie libre de Gibbs de la solution idéale dans l'équation (2.4).

$$G^M = G - G^{id} \quad (2.4)$$

Par définition, les grandeurs de mélange seraient nulles pour les solutions idéales. Elles sont définies uniquement pour les mélanges et n'ont aucune signification pour les fluides purs [34].

On cherche maintenant à développer l'expression (2.4) afin de mettre en évidence des relations avec d'autres grandeurs connues comme le potentiel chimique μ . G.N. Lewis proposa une expression du potentiel chimique en fonction l'activité a pour un composé i , à l'aide de la constante des gaz parfaits R et de la température T (equation (2.5)) [32]. La notion d'activité sera abordée en détail dans la section 2.1.5.

$$\mu_i = \mu_i^o + RT \ln a_i \quad (2.5)$$

Le potentiel chimique de référence, μ^o , peut être considéré comme le potentiel chimique du composé pur, non mélangé. Ainsi, pour aller plus loin avec la relation (2.2) citée précédemment qui met en lien le potentiel chimique avec l'énergie libre de Gibbs, on peut décomposer cette énergie comme une somme des contributions de chaque composé k , l'équation (2.6) formalise cette idée [34].

$$G = \sum_{i=1}^k n_i \mu_i \iff g = \sum_{i=1}^k x_i \mu_i \quad (2.6)$$

En combinant les équations (2.2), (2.4) et (2.5), on obtient une nouvelle relation (2.7), où l'expression de G^M devient facilement identifiable.

$$\begin{aligned} G &= G^{id} + G^M = \sum_{i=1}^k n_i \mu_i^o + \sum_{i=1}^k n_i RT \ln a_i \\ \rightarrow G^M &= \sum_{i=1}^k n_i RT \ln a_i \iff \frac{G^M}{RT} = \sum_{i=1}^k n_i \ln a_i \end{aligned} \quad (2.7)$$

Au final, il n'est pas utile de calculer l'énergie de Gibbs de la solution idéale, G^{id} , ni celle totale, G , car l'objectif est de minimiser uniquement l'énergie qui dévie de l'idéalité, G^M . C'est pourquoi cette dernière relation (2.7) est indispensable pour le calcul de flash et l'élaboration d'un diagramme de phases, elle représente la fonction objectif à minimiser [40].

2.1.5 Activité

Dans le cas de réactions chimiques, l'activité est assimilée à la constante d'équilibre associée. Pour un équilibre liquide-vapeur, elle représente une concentration "effective" du composé i . L'équation (2.8) montre cette relation avec la fraction molaire ou massique x et introduit le coefficient d'activité γ , sans dimension comme l'activité [41].

$$a_i = \gamma_i x_i \iff \gamma_i = \frac{a_i}{x_i} \quad (2.8)$$

La principale difficulté rencontrée avec la relation (2.8) est qu'elle n'est pas linéaire, car le coefficient d'activité γ varie également avec cette fraction x . En somme, on considèrera l'activité comme un écart à l'idéalité et similaire à la concentration "corrigée" ou "altérée" du composé i par le coefficient d'activité. Il faut noter également qu'il existe plusieurs valeurs numériques possibles de l'activité pour un même composé et un même état thermodynamique, elle dépendra de l'état de référence choisi. Si cette référence est une solution idéale, le coefficient d'activité sera donc égal à 1 et $a_i = x_i$. C'est le paramètre le plus difficile à estimer, un modèle mécanistique présenté dans la section 2.4 permettra de le calculer [42].

2.1.6 Fugacité

Dans le cas des équilibres liquide-vapeur, il est essentiel d'introduire la fugacité pour faire le lien entre les grandeurs des phases présentes dans le système. Elle permet de représenter les équilibres car elle est directement liée à l'énergie libre de Gibbs G . Pour un équilibre de deux phases A et B , $G_i^{M,A} = G_i^{M,B}$, ce qui implique donc $\hat{f}_i^A = \hat{f}_i^B$, où \hat{f}_i est la fugacité du composé i mélangé dans la phase. Plus concrètement, la notion de fugacité se traduit comme la tendance du composé i à "s'échapper" de la phase considérée. A l'équilibre, cette tendance est équivalente dans les deux phases (équation (2.9)) [43].

$$\hat{f}_i^{gaz} = \hat{f}_i^{liq} \quad (2.9)$$

A partir cette équivalence, on admet alors l'expression générale pour décrire les équilibres liquide-vapeur (équation (2.10)) qui met en évidence certains paramètres traités plus tôt comme le coefficient d'activité [43].

$$\hat{\Phi}_i \cdot y_i \cdot P = \gamma_i \cdot x_i \cdot \hat{\Phi}_i^{sat} \cdot P_i^{sat} \cdot \mathcal{P} \quad (2.10)$$

Brièvement, le coefficient de fugacité, $\hat{\Phi}_i$, le coefficient de fugacité à saturation, $\hat{\Phi}_i^{sat}$, et la correction de Poynting, \mathcal{P} , sont d'autres facteurs qui tiennent compte de la non idéalité de la phase gazeuse et liquide, ils sont particulièrement utiles pour des systèmes dont la pression est élevée (supérieure à 10 bars). Cependant, les systèmes qui seront étudiés ici sont ceux dont la pression est égale ou inférieure à la pression atmosphérique, il est alors possible d'appliquer tout simplement la loi des gaz parfaits pour la phase gazeuse et de simplifier certains paramètres de la partie liquide pour ne garder que son coefficient d'activité γ . Ces hypothèses impliquent $\hat{\Phi}_i = 1$, $\hat{\Phi}_i^{sat} = 1$ et $\mathcal{P} = 1$ et conduisent à l'équation (2.11) qui est, au final, la loi de Raoult modifiée [43, 44].

$$y_i \cdot P = \gamma_i \cdot x_i \cdot P_i^{sat} \quad (2.11)$$

Il faudra alors déduire entre autres les fractions molaires gazeuses de chaque composé, y_i , les fractions molaires liquides, x_i et les pressions de vapeur saturante P_i^{sat} d'un système soumis à une pression P .

Maintenant que tous les outils théoriques essentiels ont été présentés, le système considéré peut être détaillé ainsi que le modèle choisi pour la résolution des équations qui en découlent.

2.2 Calcul des équilibres liquide-vapeur du binaire méthyl tert-butyl-éther / eau

Comme présenté au début de ce chapitre, le calcul d'un équilibre liquide-vapeur commence par l'application du second principe de la thermodynamique pour une transformation se produisant à T et P constantes (équation (2.12)).

$$\delta S_C \geq 0 \Leftrightarrow dG_{T,P} \leq 0 \quad (2.12)$$

Lorsque la transformation atteint un état d'équilibre, la variation d'entropie créée est nulle ($\delta S_C = 0$), l'énergie de Gibbs atteint un minimum ($dG_{T,P} = 0$). On s'intéresse au système binaire qui implique du méthyl-ter-butyl-éther (1) et de l'eau (2) sous pression atmosphérique ($P = P^o = 1.01325$ bar). On

note z_1 et z_2 les fractions molaires globales des constituants (1) et (2). Supposons qu'à T , P et z_1 fixées, ce système binaire soit biphasique, la figure 2.2 fixe les notations employées.

<p style="text-align: center;">Phase II (liquide ou vapeur)</p> <p>Quantité de matière : $n_{II} = n_{1,II} + n_{2,II}$</p> <p>Fractions molaires : $x_{2,II} = 1 - x_{1,II}$ $x_{1,II} = n_{1,II}/n_{II}$</p> <p>Proportion molaire de phase II : $\%_{II} = n_{II}/n_T$</p>	}	<p>Grandeurs globales du système :</p> <p>Température : T</p> <p>Pression : P</p> <p>Quantité de matière globale n_T : $n_T = n_I + n_{II} = n_1 + n_2$</p> <p>Fraction molaire globale z_1 : $z_1 = n_1/n_T = \%_I \cdot x_{1,I} + \%_{II} \cdot x_{1,II}$</p>
<p style="text-align: center;">Phase I (liquide ou vapeur)</p> <p>Quantité de matière : $n_I = n_{1,I} + n_{2,I}$</p> <p>Fractions molaires : $x_{2,I} = 1 - x_{1,I}$ $x_{1,I} = n_{1,I}/n_I$</p> <p>Proportion molaire de phase I : $\%_I = n_I/n_T$</p>		

Ici, $x_{1,I}$ et $x_{1,II}$ désignent les fractions molaires en constituant (1) dans les phases I et II . Les fractions molaires en constituant (2) dans chaque phase sont alors déduites par différence avec le constituant (1) selon l'expression (2.13).

$$x_{2,I} = 1 - x_{1,I} \quad \text{et} \quad x_{2,II} = 1 - x_{1,II} \quad (2.13)$$

En reprenant l'expression (2.7), on définit à présent l'énergie de Gibbs molaire totale de mélange du système binaire (équation (2.14)).

$$\frac{g^M}{RT} = \%_I \cdot \frac{g^{M,I}(x_{1,I})}{RT} + \%_{II} \cdot \frac{g^{M,II}(x_{1,II})}{RT} \quad (2.14)$$

Cette énergie est composée de $g^{M,I}(x_{1,I})$ et $g^{M,II}(x_{1,II})$ les énergies de Gibbs molaires totales de mélange des phases I et II proportionnée par les quantités $\%_I$ et $\%_{II}$, proportions molaires des phases I et II . Ces dernières sont reliées aux fractions molaires par les relations en (2.15).

$$\begin{cases} \%_{II} = \frac{z_1 - x_{1,I}}{x_{1,II} - x_{1,I}} \\ \%_I = 1 - \%_{II} \end{cases} \quad (2.15)$$

Pour calculer un équilibre biphasique, il suffit de minimiser la fonction $\frac{g^M}{RT}$ par rapport aux deux variables $x_{1,I}$ et $x_{1,II}$. Si à l'issue de la procédure, le minimum global de la fonction est obtenu avec $x_{1,I} = x_{1,II}$, le mélange est déclaré monophasique. Au contraire, si l'on obtient $x_{1,I} \neq x_{1,II}$, le mélange est déclaré biphasique.

En remplaçant (2.15) dans (2.14), on aboutit à un système d'équations (2.16), (2.17) et (2.18) qui seront intégrées dans le programme d'optimisation.

$$\frac{g^M}{RT} = \frac{x_{1,II} - z_1}{x_{1,II} - x_{1,I}} \cdot \frac{g^{M,I}(x_{1,I})}{RT} + \frac{z_1 - x_{1,I}}{x_{1,II} - x_{1,I}} \cdot \frac{g^{M,II}(x_{1,II})}{RT} \quad (2.16)$$

$$\begin{cases} \frac{g^{M,I}}{RT} = \min \left\{ \frac{g^{M,I_{liq}}}{RT}, \frac{g^{M,I_{gaz}}}{RT} \right\} \\ \frac{g^{M,I_{liq}}}{RT} = x_{1,I} \cdot \ln a_{1,I}^{liq} + (1 - x_{1,I}) \cdot \ln a_{2,I}^{liq} \\ \frac{g^{M,I_{gaz}}}{RT} = x_{1,I} \cdot \ln a_{1,I}^{gaz} + (1 - x_{1,I}) \cdot \ln a_{2,I}^{gaz} \end{cases} \quad (2.17)$$

$$\begin{cases} \frac{g^{M,II}}{RT} = \min \left\{ \frac{g^{M,II_{liq}}}{RT}, \frac{g^{M,II_{gaz}}}{RT} \right\} \\ \frac{g^{M,II_{liq}}}{RT} = x_{1,II} \cdot \ln a_{1,II}^{liq} + (1 - x_{1,II}) \cdot \ln a_{2,II}^{liq} \\ \frac{g^{M,II_{gaz}}}{RT} = x_{1,II} \cdot \ln a_{1,II}^{gaz} + (1 - x_{1,II}) \cdot \ln a_{2,II}^{gaz} \end{cases} \quad (2.18)$$

Pour un mélange binaire comme décrit sur la figure 2.2, si l'on applique les équations (2.8) et (2.11) la résolution des activités pour chaque phase dépendra alors de la pression P appliquée :

- Pour le constituant (1) :
 - Si $P > P_1^{sat}(T)$:
 - * Phase I (liquide ou gaz) :

$$\begin{cases} a_{1,I}^{liq} = x_{1,I} \cdot \gamma_1(T, x_{1,I}) \\ a_{1,I}^{gaz} = \frac{P \cdot x_{1,I}}{P_1^{sat}(T)} \end{cases} \quad (2.19)$$

- * Phase II (liquide ou gaz) :

$$\begin{cases} a_{1,II}^{liq} = x_{1,II} \cdot \gamma_1(T, x_{1,II}) \\ a_{1,II}^{gaz} = \frac{P \cdot x_{1,II}}{P_1^{sat}(T)} \end{cases} \quad (2.20)$$

- Si $P < P_1^{sat}(T)$:
- * Phase I (liquide ou gaz) :

$$\begin{cases} a_{1,I}^{liq} = \frac{P_1^{sat}(T) \cdot x_{1,I} \cdot \gamma_1(T, x_{1,I})}{P} \\ a_{1,I}^{gaz} = x_{1,I} \end{cases} \quad (2.21)$$

- * Phase II (liquide ou gaz) :

$$\begin{cases} a_{1,II}^{liq} = \frac{P_1^{sat}(T) \cdot x_{1,II} \cdot \gamma_1(T, x_{1,II})}{P} \\ a_{1,II}^{gaz} = x_{1,II} \end{cases} \quad (2.22)$$

- De la même manière, pour le constituant (2) :
 - Si $P > P_2^{sat}(T)$:
 - * Phase I (liquide ou gaz) :

$$\begin{cases} a_{2,I}^{liq} = (1 - x_{1,I}) \cdot \gamma_2(T, x_{1,I}) \\ a_{2,I}^{gaz} = \frac{P \cdot (1 - x_{1,I})}{P_2^{sat}(T)} \end{cases} \quad (2.23)$$

* Phase II (liquide ou gaz) :

$$\begin{cases} a_{2,II}^{liq} = (1 - x_{1,II}) \cdot \gamma_2(T, x_{1,II}) \\ a_{2,II}^{gaz} = \frac{P \cdot (1 - x_{1,II})}{P_2^{sat}(T)} \end{cases} \quad (2.24)$$

– Si $P < P_2^{sat}(T)$

* Phase I (liquide ou gaz) :

$$\begin{cases} a_{2,I}^{liq} = \frac{P_2^{sat}(T) \cdot (1 - x_{1,I}) \cdot \gamma_2(T, x_{1,I})}{P} \\ a_{2,I}^{gaz} = (1 - x_{1,I}) \end{cases} \quad (2.25)$$

* Phase II (liquide ou gaz) :

$$\begin{cases} a_{2,II}^{liq} = \frac{P_2^{sat}(T) \cdot (1 - x_{1,II}) \cdot \gamma_2(T, x_{1,II})}{P} \\ a_{2,II}^{gaz} = (1 - x_{1,II}) \end{cases} \quad (2.26)$$

Pour calculer ces activités, deux variables doivent être encore exprimées : la pression de vapeur saturante P^{sat} et le coefficient d'activité γ .

2.3 Calcul des pressions de vapeur saturante

Afin de pouvoir calculer les pressions de vapeur saturante, P_i^{sat} , d'un composé i pur, l'équation d'Antoine non étendue est utilisée. Trois paramètres A , B et C sont obtenus par regression sur base de données expérimentales disponible dans la littérature pour donner une fonction de pression de vapeur saturante en fonction de la température, T , pour chaque composé i . Les pressions de vapeur saturante des constituants (1) et (2) sont données par l'équation d'Antoine non étendue (2.27) [45, 46].

$$\log(P_i^{sat}) = A_i - \frac{B_i}{T + C_i} \quad (2.27)$$

Avec P_i^{sat} en bar et T en degrés celsius, les coefficients correspondants pour les deux composés sont résumés dans le tableau 2.1.

i	Composé	A_i	B_i	C_i	T_{min} (°C)	T_{max} (°C)
1	Methyl-tert-butyl-ether	3.9698	1097.67	221.8468	-108	224
2	Eau	5.1792	1722.09	232.8847	0	374

Tableau 2.1: Coefficients d'Antoine pour le méthyl-ter-butyl-éter et l'eau, le domaine de validité des coefficients est défini entre T_{min} et T_{max} [46]

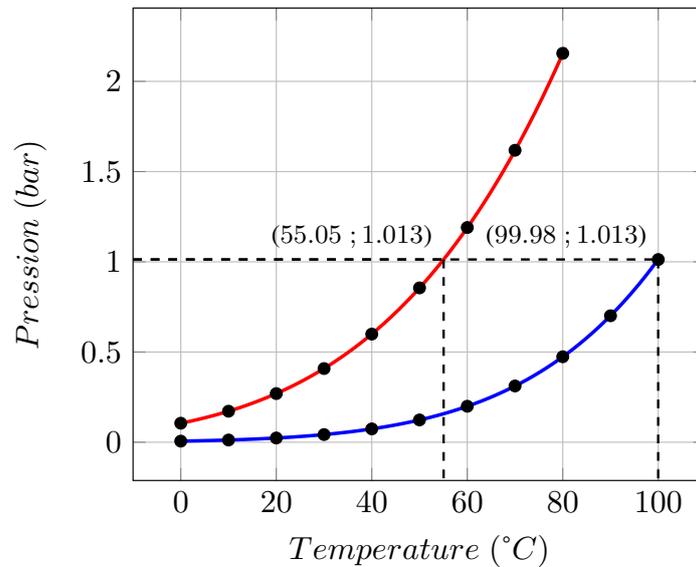


Figure 2.3: Courbe de vapeur saturante du methyl-ter-butyl-ether entre 0°C et 80°C (en rouge) et de l'eau entre 0°C et 100°C (en bleu) donnés par l'équation d'Antoine

Sur la figure 2.3, on retrouve bien les températures d'ébullition des corps purs à pression atmosphérique, ces valeurs particulières sont visibles sur un diagramme de phases pour $x_1 = 1$ et $x_2 = 1$. Concrètement, ces courbes sont utiles car elles fournissent pour chaque température donnée, les pressions de vapeur saturantes des deux composés qui sont introduites dans les équations (2.19) à (2.26).

2.4 Modèle NRTL

L'interpolation et l'extrapolation des données thermodynamiques des mélanges liquides sont indispensables pour l'ingénierie chimique. Le modèle des solutions idéales est utile si l'on souhaite obtenir une première approximation, mais les écarts à l'idéalité sont souvent importants et méritent d'être quantifiés. Ces écarts sont exprimés par les fonctions d'excès qui dépendent de la concentration des composés, de la température et de la pression. Wilson démontra que l'excès d'énergie de Gibbs peut être exprimée en fonction des compositions locales, c'est à dire l'agencement des molécules entre elles et, ceci, pour une large variété de composés. Ce concept n'est pas très rigoureux dans l'absolu, certaines hypothèses sont nécessaires, parfois incohérentes voire empiriques selon les cas. Cependant, il a démontré son efficacité depuis plus de trente ans [47, 48]. A partir des équations de Wilson, d'autres modèles plus spécifiques ont vu le jour, dont les modèles NRTL et UNIQUAC qui sont aujourd'hui des modèles très utilisés dans l'industrie pharmaceutique [49]. La figure 2.4 est un exemple d'arbre décisionnel qui, en fonction des conditions et de la nature des composés, permet d'appliquer le modèle le plus adéquat au problème considéré. Classiquement, on retrouvera par exemple les modèles de Peng-Robinson, Redlich-Kwong-Soave ou Lee-Kesler dans l'industrie pétrochimique, les composés traités étant principalement non polaires (hydrocarbures) et sous haute pression [50].

Ici, on préférera s'intéresser au modèle NRTL, développé par Renon et Prausnitz en 1968, qui s'applique plus spécifiquement à des solutions non électrolytiques, polaires et à basses pressions [47]. L'équilibre liquide-liquide ne sera pas étudié pour ce système. On peut représenter schématiquement les interactions moléculaires comme présenté sur la figure 2.5. Pour un mélange binaire, deux cas de figure

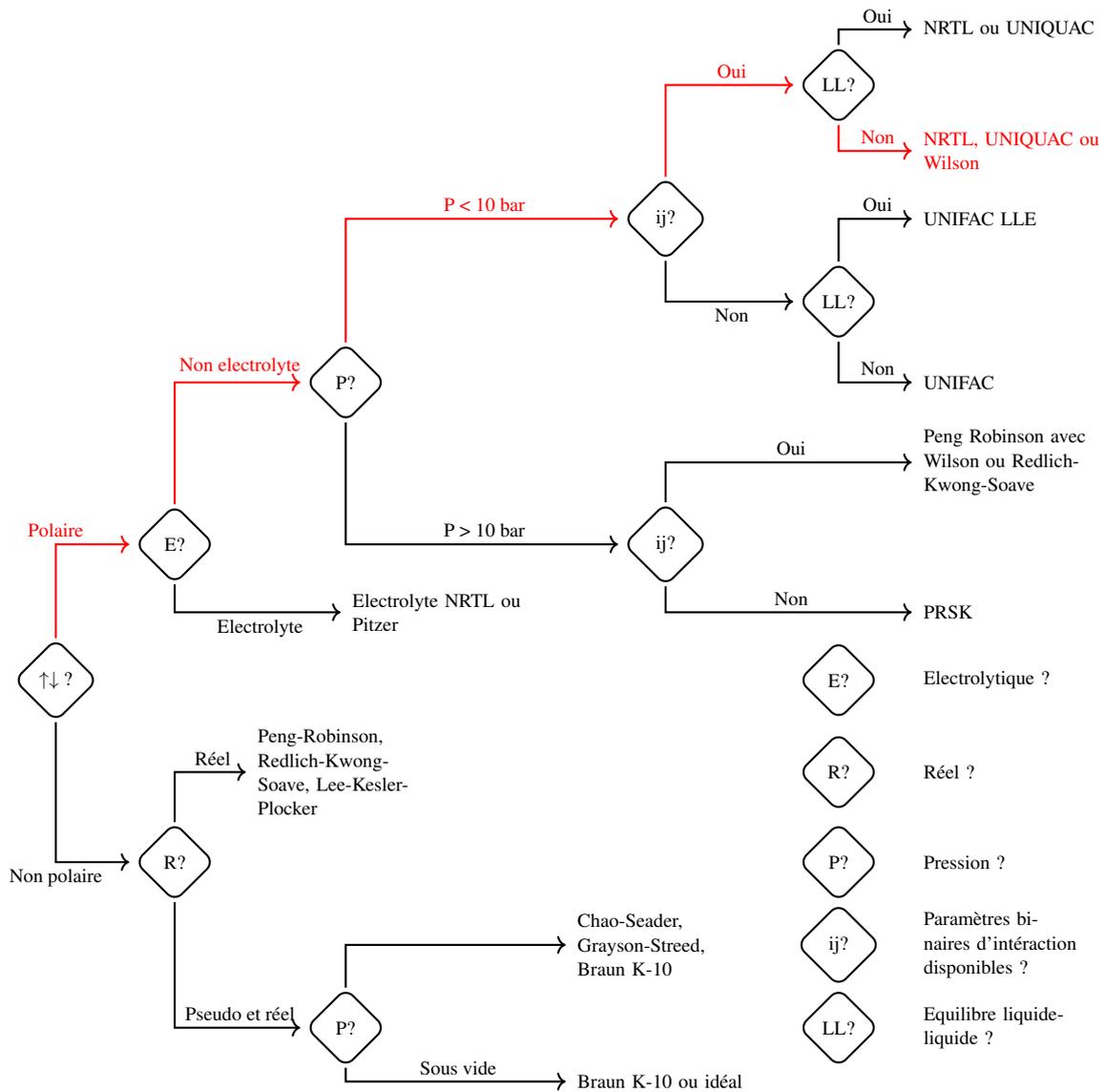


Figure 2.4: Arbre décisionnel pour le choix du modèle à employer selon la solution réelle considérée. En rouge le chemin suivi pour l'exemple méthyl tert-butyl-éther - eau [50]

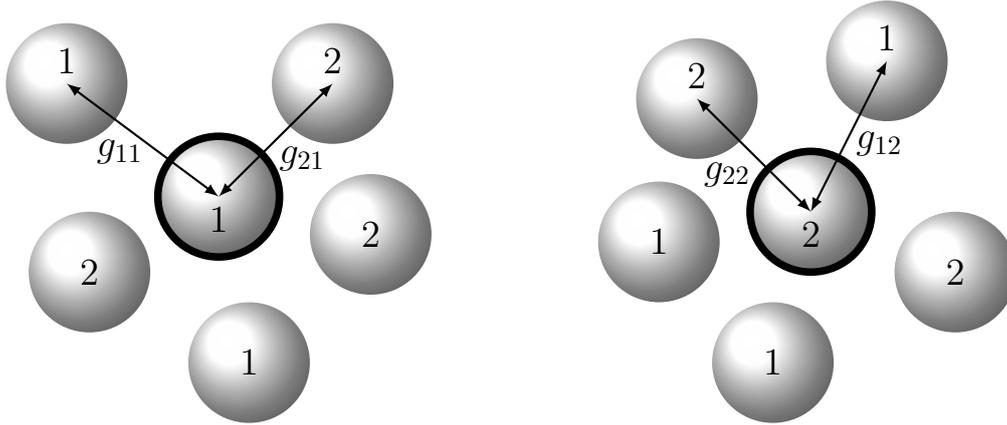


Figure 2.5: Schéma représentatif des interactions prises en compte dans le modèle NRTL d'un mélange binaire (molécules 1 et 2)

sont possibles. En considérant une molécule 1 au centre, deux types d'interactions sont possibles, soit avec une autre molécule 1, g_{11} , soit avec une molécule 2, g_{21} . Ce même raisonnement s'applique avec la molécule 2 au centre. Il est important de noter ces deux cas de figure, car en réalité g_{12} n'est pas égal à g_{21} , le système est dit "non symétrique" [45].

L'expression des logarithmes des coefficients d'activité (γ_1 et γ_2) des constituants en phase liquide proposé par Renon et Prausnitz est décrite dans le système (2.28) dans lequel trois inconnues apparaissent : g_{12} , g_{21} et α . Ils constituent les paramètres caractéristiques d'un système pour le modèle NRTL.

$$\left\{ \begin{array}{l} \ln \gamma_1(T, x_1) = x_2^2 \left[\tau_{21} \left(\frac{G_{21}}{x_1 + x_2 G_{21}} \right)^2 + \frac{\tau_{12} G_{12}}{(x_2 + x_1 G_{12})^2} \right] \\ \ln \gamma_2(T, x_1) = x_1^2 \left[\tau_{12} \left(\frac{G_{12}}{x_2 + x_1 G_{12}} \right)^2 + \frac{\tau_{21} G_{21}}{(x_1 + x_2 G_{21})^2} \right] \\ x_2 = 1 - x_1 \\ \tau_{12} = \frac{g_{12}}{RT} \\ \tau_{21} = \frac{g_{21}}{RT} \\ G_{12} = \exp(-\alpha \tau_{12}) \\ G_{21} = \exp(-\alpha \tau_{21}) \end{array} \right. \quad (2.28)$$

L'application de ce type de modèle impose un énorme travail de corrélation, car un système de dix constituants, plutôt commun dans la pratique, contient quarante-cinq systèmes binaires [51]. Fort heureusement, des bases de données existent et proposent les valeurs des paramètres pour beaucoup de systèmes binaires usuels. On peut citer la plus connue et la plus riche qui a été initiée par l'Université de Dortmund en Allemagne : la DDB ("Dortmund Data Bank"). Les valeurs du système méthyl tert-butyl-éther/eau sont présentées en (2.29) et proviennent de cette banque [52].

$$\left\{ \begin{array}{l} g_{12} = 5704.88 \quad J \cdot mol^{-1} \\ g_{21} = 9203.89 \quad J \cdot mol^{-1} \\ \alpha = 0.3 \end{array} \right. \quad (2.29)$$

Un autre paramètre, α , est introduit, il représente le "caractère non aléatoire" d'un mélange. En d'autres termes, pour un liquide dont la distribution locale est aléatoire autour de la molécule centrale, α sera nul. En pratique, α prendra le plus souvent des valeurs discrètes comme 0.2, 0.3 ou bien 0.48. Cette dernière est le plus souvent attribuée à des systèmes aqueux, car une valeur élevée reflète une structure ordonnée par des liaisons hydrogène. Afin d'accroître la précision du modèle, il est possible d'utiliser deux paramètres plutôt qu'un seul : α_{12} et α_{21} , représentatifs également de chaque composé. Ce type de modèle étant moins courant, un seul paramètre sera utilisé ici.

Il arrive cependant qu'aucune donnée expérimentale ne soit disponible pour le système en question. Dans ce cas, il faut savoir les générer au laboratoire ou bien considérer une autre méthode, bien plus théorique, basée sur les "contributions de groupes". Brièvement, cette méthode prédit en quelque sorte le comportement des interactions moléculaires entre deux espèces d'un système en fonction de leurs groupes fonctionnels (éther, alcane, cétones, ou encore aldéhydes). Le nombre de groupements atomiques étant moins vaste que le nombre d'espèces chimiques, la quantité de paramètres à corrélérer est donc drastiquement réduite. Le problème de cette méthode est qu'elle suppose, par exemple, qu'un groupement méthyl final d'une paraffine est identique à celui compris dans une molécule d'éthanol. L'environnement n'est donc plus pris en compte. Ainsi est né le modèle UNIFAC, bien moins puissant que les modèles précédemment cités mais qui a son importance si les données des autres modèles viennent à manquer [51].

2.5 Construction du diagramme des équilibres liquide-vapeur

2.5.1 Types de diagrammes d'équilibres liquide - vapeur

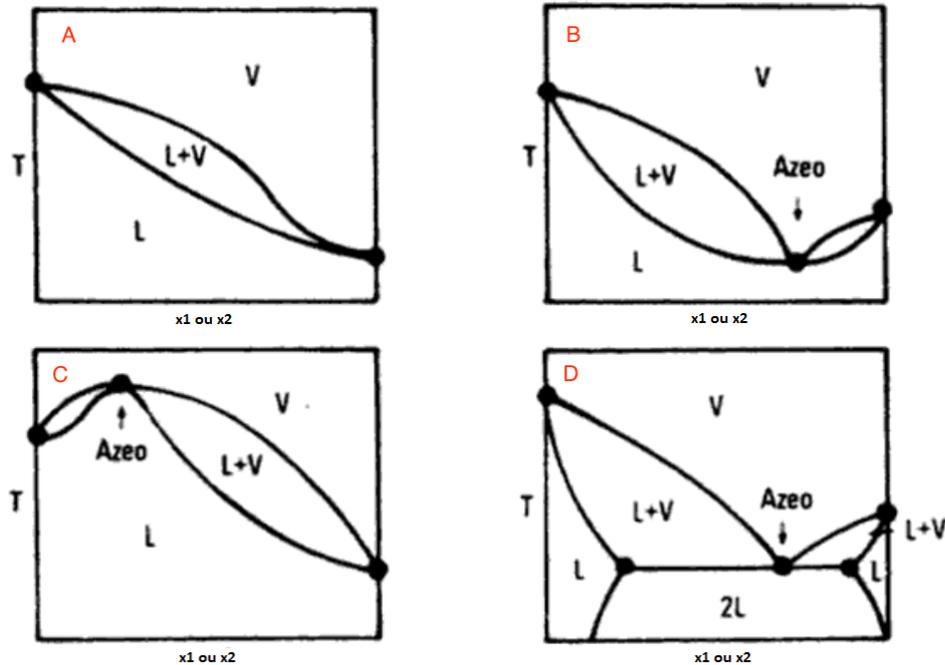


Figure 2.6: Schémas des différents diagrammes liquide-vapeur usuels rencontrés pour des systèmes binaires à pression constante en fonction de la fraction molaire ou massique de l'un des deux constituants et de la température [53]

La construction d'un diagramme des équilibres liquide-vapeur est très utile pour le développement des procédés de distillation, il renseigne directement sur les compositions de chaque phase pour une température et pression fixées. L'exploitation de ces diagrammes permet d'effectuer certains calculs pour estimer les quantités de solvants à distiller pour atteindre les compositions ciblées.

La figure 2.6 présente plusieurs exemple-types de diagrammes, dont le premier, A, qui est un équilibre classique de type méthanol/eau. Les deux seuls point particuliers sont les températures d'ébullition des deux composés aux extrémités. Une composition globale comprise dans le domaine liquide-vapeur (L+V) verra la composition de sa phase liquide différente de sa phase vapeur. Elles sont données par l'intersection horizontale des domaines liquide et vapeur avec la température du système.

Cependant, beaucoup de mélanges montrent un comportement azéotropique pour certaines compositions, de telle sorte que la phase liquide en équilibre avec sa phase vapeur ont la même composition. Les trois autres exemples de la figure 2.6 montrent ce point particulier qui peut être positif (B et D) ou bien négatif (C). Dans les procédés de séparation par distillation, ce point est absolument à éviter si l'on souhaite isoler un composé pur. Un exemple classique de l'industrie chimique concerne l'éthanol à 96% présentant une composition azéotropique avec l'eau (4%), similaire au diagramme B sur la figure 2.6. L'obtention de l'éthanol absolu n'est pas possible en l'état, car distiller un tel mélange ne peut pas con-

centrer l'éthanol, la phase vapeur en équilibre étant de même composition. L'astuce est donc d'ajouter du benzène pour modifier ce comportement, de sorte à "sauter" le point azéotropique, et distiller le nouveau mélange [53].

Le dernier diagramme, D, de la figure 2.6 est particulier dans le sens où, il présente un azéotropisme positif qui rejoint le domaine liquide-liquide, ce comportement est appelé "azéotropisme hétérogène", comme pour le mélange butanol/eau. A la composition de l'azéotrope, deux phases liquides co-existent, ce n'est seulement que pour des compositions très faibles d'un composé dans l'autre, et vice-versa, qu'une seule phase liquide existe.

Le but est maintenant de déterminer le type de comportement lié au système méthyl tert-butyl-éther/eau. La génération de tels diagrammes est essentielle pour adopter la bonne stratégie de séparation du mélange binaire.

2.5.2 Résultats pour le système méthyl tert-butyl-éther/eau

Le programme en annexe 2 permet de générer le diagramme des équilibres liquide-vapeur pour le système méthyl tert-butyl-éther/eau sur base de tout ce qui a été développé dans ce chapitre (figure 2.7). La technique de construction itérative est la suivante :

1. Température de départ : point d'ébullition du corps pur
2. Minimisation de l'énergie de Gibbs pour la température donnée via l'algorithme
3. Collecte des résultats sur les compositions molaires dans chaque phase
4. Diminution de la température selon un pas défini
5. Retour à l'étape 2 jusqu'à atteindre la température la plus basse du diagramme, ici 52.04°C
6. Arrêt du programme et compilation des résultats

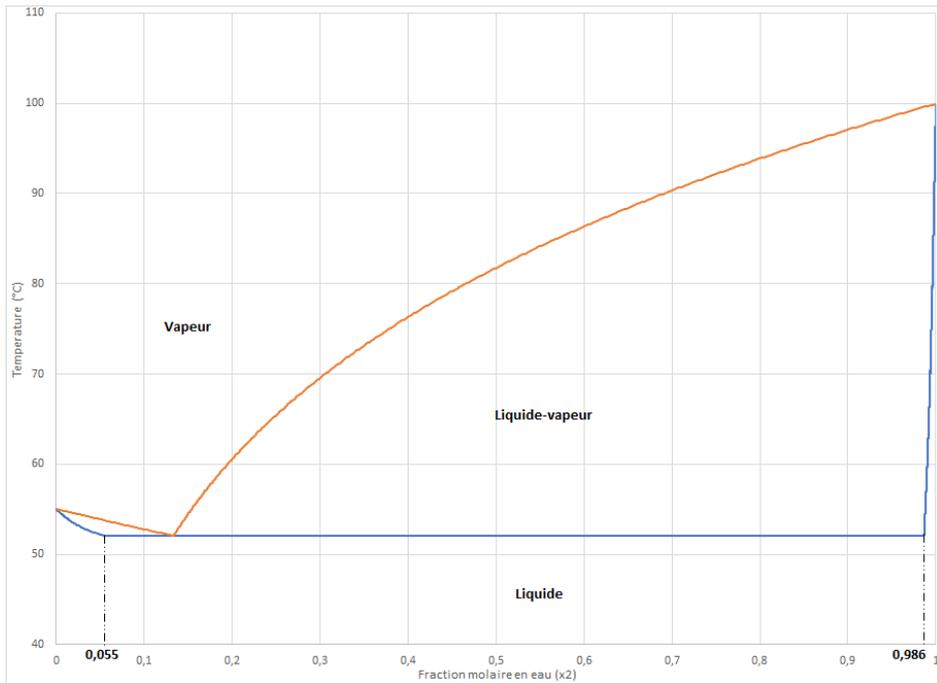


Figure 2.7: Diagramme (x,T) des équilibres liquide-vapeur du mélange méthyl tert-butyl-éther/eau à pression atmosphérique généré par l'algorithme Krill Herd, en bleu la composition liquide, en orange la composition vapeur

Premièrement, les températures d'ébullition de l'eau pour $x = 1$ et du méthyl tert-butyl-éther pour $x = 0$ à pression atmosphérique, données par la figure 2.3, sont bien retrouvées. Deuxièmement, le mélange méthyl tert-butyl-éther/eau montre un azéotropisme hétérogène. La composition azéotropique, à pression atmosphérique, est donnée pour un mélange de 13.3% molaire en eau. Et la séparation liquide-liquide à 52.04°C donne également deux informations :

- le méthyl tert-butyl-éther ne peut contenir plus de 5.5% molaire en eau
- la méthyl tert-butyl-éther est miscible dans l'eau à hauteur de 1.4% molaire ($1 - 0,986$)

A titre d'exemple, la figure 2.8 représente la courbe de l'énergie de Gibbs de mélange à la température de 52.04°C, on voit graphiquement que les deux minima de la fonction correspondante correspondent bien aux deux compositions citées précédemment.

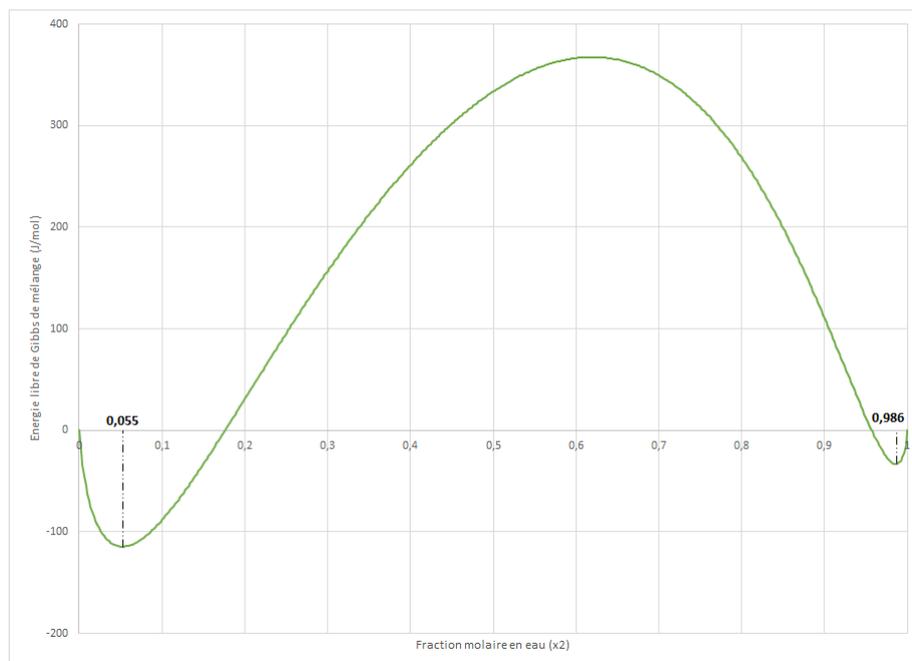


Figure 2.8: Energie libre de Gibbs de mélange du système méthyl tert-butyl-éter/eau en fonction de la fraction molaire en eau à 52.04°C et à pression atmosphérique

Afin de déterminer la pertinence de ces résultats, un diagramme similaire a été généré également avec le programme *ASPEN V9*, référence pour l'industrie pétrochimique, qui utilise un algorithme de minimisation de type quasi-Newton appelé "Boston-Britt" [54]. L'erreur absolue calculée par une simple méthode des trapèzes, soit la différence d'aire entre les deux courbes générées, est alors de 0.02%. Cet écart est tout à fait acceptable et démontre bien l'efficacité du bioalgorithme pour la minimisation de l'énergie de Gibbs de mélange ainsi développé.

Dans le chapitre 3, on s'intéressera à l'emploi de ce diagramme pour aboutir à une procédure de distillation à l'échelle pilote. Le procédé de fabrication d'un médicament anti-épileptique en phase 3 actuellement en développement à UCB Pharma en Belgique, met en jeu le principe actif issu d'une dernière réaction, dissous dans le méthyl tert-butyl-éter. Suite à de multiples extractions aqueuses en amont, l'opération de distillation en question vise à éliminer l'eau résiduelle dans la phase organique pour obtenir une solution quasi anhydre. On étudiera alors la quantité de solvant à employer pour atteindre cet objectif à l'aide des outils développés. Des procédures alternatives seront envisagées pour réduire cette consommation.

3 Application à un cas concret de distillation industrielle

3.1 Objectifs de la distillation industrielle

Dans les usines chimiques conventionnelles, trois grandes étapes peuvent être retrouvées : la préparation, la réaction et l'étape de séparation. Entre 60 et 80% des coûts sont liés aux étapes de séparation, dont la distillation [55]. La distillation est le procédé le plus ancien et le plus universel parmi les technologies chimiques et les autres branches de l'industrie comprenant des séparations de mélanges. Elle reste avantageuse car relativement efficace par rapport aux procédés d'extraction, de cristallisation et de séparation membranaire [56]. C'est une opération dans laquelle les composés d'un mélange liquide sont séparés par vaporisation. La phase vapeur est normalement plus riche en espèces les plus volatiles comparativement à la phase liquide avec laquelle elle est en équilibre. Les systèmes azéotropiques ne peuvent pas être séparés par distillation, cependant dans certains cas le point azéotrope peut être modifié en changeant les conditions opératoires. Les problèmes de séparation sont aussi retrouvés avec les mélanges de liquides ayant des volatilités similaires ; dans ce cas elle peut être facilitée en ajoutant un troisième composé [55].

Les procédés de distillation peuvent être aussi potentiellement critiques pour les solutés, ils impliquent des températures élevées qui peuvent favoriser des cinétiques de décomposition. Généralement, une étude sur la stabilité thermique des composés en solution est d'abord établie afin de juger des conditions à appliquer pour les premiers essais. De plus, certains d'entre eux peuvent modifier les équilibres liquide-vapeur de façon plus ou moins significative [57].

Il existe une grande variété de méthodes de distillation mais elles peuvent être regroupées en deux catégories. Dans l'une, la vapeur produite par ébullition de la phase liquide est chassée puis condensée en dehors du réacteur, elle peut être de type "simple" ou "flash". Dans la seconde méthode, une partie de la vapeur condensée retourne dans le réacteur de telle sorte que le liquide est à contre courant du flux gazeux, une telle méthode est dite de "fractionnement" [58].

3.1.1 Distillation flash

La distillation flash, ou distillation à l'équilibre, est un procédé où un liquide saturé est soumis à une réduction de pression, le liquide se vaporise alors en partie donnant un équilibre liquide-vapeur. Cette opération est le plus souvent continue, un fluide est transporté via une conduite jusqu'à une cuve, dite "flash", produisant en continu, une phase vapeur et une phase liquide (figure 3.2). Le fluide d'entrée peut être un liquide ou de la vapeur. Pour un système binaire, la phase vapeur aura alors une composition différente de la phase liquide en accordance avec le diagramme de phases correspondant pour une température et une pression données (figure 3.1). A l'inverse d'une distillation fractionnée, il n'existe qu'un seul "étage" de séparation, c'est à dire un seul équilibre liquide-vapeur dans la cuve.

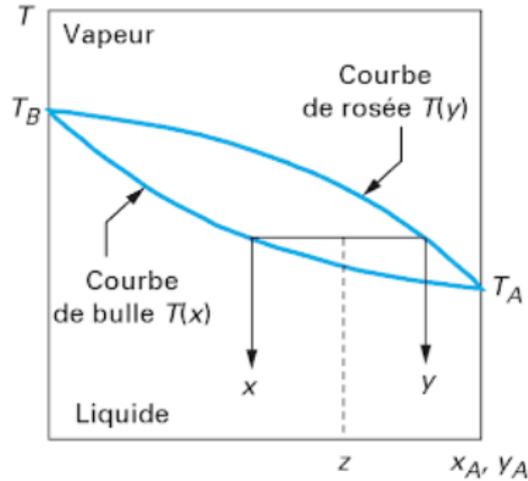


Figure 3.1: Schéma exemple d'un diagramme des équilibres liquide-vapeur d'un binaire A et B , la température T est fonction de la composition en composé A dans le mélange [59]

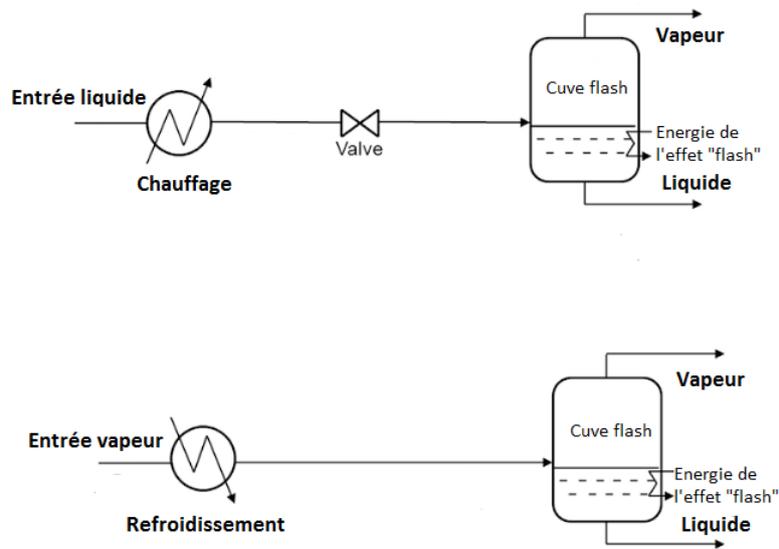


Figure 3.2: Schéma général du procédé de la distillation flash pour deux cas : entrée liquide ou vapeur [60]

Ce type de séparation est peu employée dans l'industrie pharmaceutique, on préférera des méthodes donnant des séparations plus "tranchées" et discontinues.

La distillation flash est typiquement retrouvée dans les systèmes réfrigérants conventionnels, par exemple les climatiseurs, où l'expansion du liquide saturé entraîne une chute de la température des deux phases. Un autre exemple d'application est l'extraction des gaz à partir du pétrole brut.

3.1.2 Distillation simple

Le procédé de distillation simple est connu depuis des siècles, les écrits d'Aristote en font déjà mention : "la vapeur d'eau de mer ne contient pas de salure, l'eau de mer est plus lourde que l'eau douce" [61]. Au Moyen Age, les moines et les docteurs ont affiné le procédé de distillation pour la production d'infusions curatives et extraits végétaux car ceux-ci sont à base d'alcool. Aujourd'hui on la retrouve sous différentes formes dans la production de parfum, liqueurs, huiles destinées à la consommation et bien évidemment, la pétrochimie. Dans le cas de la production de sel et d'eau douce à partir de l'eau salée, il n'y a qu'un seul "étage" de distillation. C'est ce qui caractérise la distillation simple et flash en opposition avec la distillation fractionnée qui, elle en a plusieurs [62].

Dans cette méthode, bien représentative du concept de la distillation, la vapeur est obtenue à partir d'un liquide porté à ébullition par l'apport d'une source de chaleur externe. Elle est entraînée en continu vers un condenseur afin de maintenir la pression constante durant toute l'opération. De la même manière que la distillation flash, les phases vapeur et liquide sont en équilibre thermodynamique comme présenté sur la figure 3.1.

En revanche, les applications sont différentes de celles vues pour la distillation flash, on cherchera davantage à concentrer ou purifier un mélange via ce procédé. C'est une opération très flexible car elle peut s'effectuer avec un réacteur usuel comparativement à une distillation fractionnée qui implique une colonne qui doit être dimensionnée en fonction du procédé considéré.

Le cas présenté dans ce document dans la partie 3.2 traitera plus en détail la distillation simple puisqu'elle est employée dans le procédé de synthèse de la molécule anti-épileptique développée à UCB Pharma.

3.1.3 Distillation fractionnée

En pratique, au lieu de réaliser une série de distillations simples pour séparer les composés volatiles d'un mélange, il est possible d'effectuer une distillation fractionnée comme présenté par la figure 3.3. Les vapeurs formées se condensent dans le garnissage de la colonne qui est plus froide que le réacteur contenant le liquide initial. Cette condensation conduit à un liquide qui sera localement plus riche en produit le plus volatil, comme décrit pour la distillation simple ou flash. Ce phénomène s'étend sur toute la colonne créant alors un gradient de concentration : en bas la composition initiale du mélange où la concentration en composé volatil est la plus basse et en haut où elle est la plus forte, voire pure dans certains cas. La température varie également comme un gradient, elle est plus basse en haut de colonne qu'à sa base. Par conséquent, le liquide dans le réacteur se concentre en composé le moins volatil et sa température augmente au cours de l'opération [63].

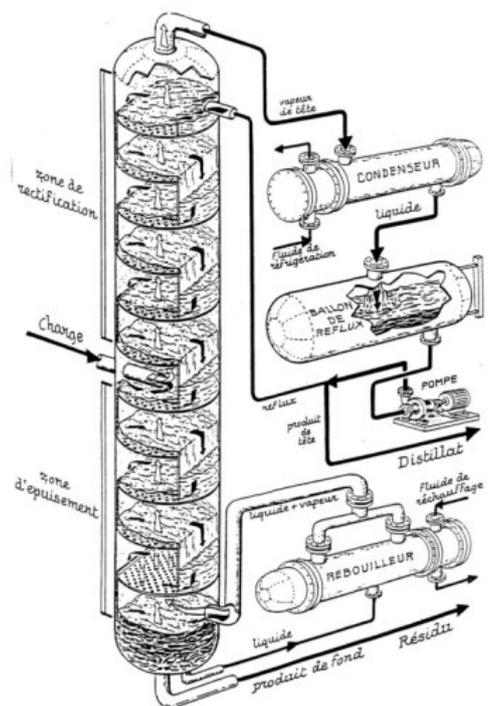


Figure 3.3: Schéma d'une distillation fractionnée avec une colonne à plateaux [64]

L'application majeure de ce procédé concerne la séparation des composés du pétrole brut. Par exemple, au premier étage il peut y avoir le point d'addition du pétrole brut qui sera chauffé. A une certaine hauteur dans la colonne, des étages physiques sont présents et les condensats y sont récoltés. Ainsi il est possible de récolter différents composés plus ou moins purs sur plusieurs étages de la colonne selon les volatilités respectives, créant des "fractions" qui seront retraitées par la suite dans d'autres colonnes à fractionnement afin d'obtenir une séparation complète du brut. Dans l'ordre du plus volatile au moins volatile, donc de haut en bas de la colonne, il y aura d'abord les gaz naturels (propane, pentane...), ensuite les essences à usage automobile, puis le diesel, les huiles et enfin les produits à usage routier tel que l'asphalte [51].

La mise en place de tels procédés est également retrouvé dans l'industrie pharmaceutique mais de façon moins courante. Les dimensions de la colonne sont directement proportionnelles à l'efficacité de la séparation souhaitée et demandent un important travail de dimensionnement qui est variable selon l'application. C'est pourquoi la distillation simple reste donc la méthode la plus représentée surtout dans la phase de développement où les changements au niveau procédé sont fréquents [58].

3.2 Cas pratique : distillation à échelle pilote dans la dernière étape de synthèse d'un principe actif

Un nouvel anti-épileptique actuellement en phase 3 à UCB Pharma est devenu l'une des priorités de développement pour l'entreprise. Sa synthèse est relativement longue et complexe, puisqu'elle consiste en douze étapes avec, pour la plupart d'entre elles, une phase de préparation, une ou plusieurs réactions chimiques avec traitement et une phase d'isolement sous forme solide ou liquide.

La dernière étape est bien évidemment la plus critique du procédé, les spécifications imposées étant strictes sur la qualité du principe actif. La figure 3.4 présente le procédé employé lors de cette étape. Après la réaction, une extraction liquide-liquide du produit est effectuée avec du méthyl tert-butyl-éther, la phase organique obtenue est alors lavée avec de l'eau pour extraire un maximum d'impuretés issues de la réaction. C'est à ce moment précis que la distillation simple intervient pour éliminer l'eau résiduelle provenant des lavages précédents. La cristallisation est ensuite engagée, et permet d'abord d'isoler le principe actif sous forme solide et également de le séparer physiquement de ses impuretés de réaction, elle est appelée "cristallisation sélective". Les étapes suivantes de filtration, lavages, et séchage contribuent aussi à la purge des impuretés et des solvants résiduels.

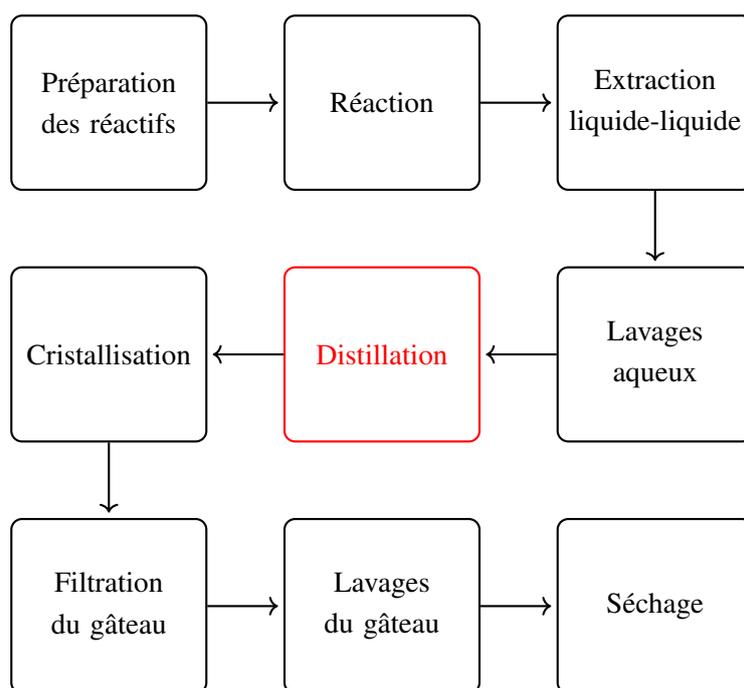


Figure 3.4: Diagramme explicatif du procédé de la dernière étape de synthèse du principe actif

Cette étape de distillation est très importante, car elle conditionne le comportement de la cristallisation en aval. En effet, avec une quantité d'eau trop élevée, la cristallisation peut se déclencher plus tôt que prévu, c'est à dire à une température plus élevée. Elle n'est donc plus contrôlée et peut entraîner une co-cristallisation des impuretés avec le principe actif où les étapes de lavages et de séchage seraient inefficaces pour les éliminer.

Ainsi, le contrôle de la quantité d'eau après distillation est essentiel pour assurer la qualité du principe actif en fin de procédé. Les limites acceptables en eau ont été largement étudiées lors de la phase de développement de ce principe actif et sont maintenant bien définies. Cependant, la spécification imposée est basse et implique un grand volume de méthyl tert-butyl-éther "frais" pour atteindre cet objectif.

Avec une volonté de réduire l'impact écologique des procédés pharmaceutiques, la démarche de minimiser les volumes de solvants employés dans un procédé de synthèse est initiée dès la phase de développement. On étudiera dans un premier temps l'application du diagramme avec la procédure actuelle de distillation et la pertinence du modèle comparativement aux résultats à l'échelle pilote. Dans un deuxième temps, une procédure de distillation continue sera proposée et comparée avec la distillation

batch actuelle. Enfin, certains points à améliorer pour la modélisation seront développés.

3.2.1 Objectifs

A la suite des lavages aqueux, la phase organique est saturée en eau. Selon le diagramme des équilibres liquide-vapeur présenté en 3.5, la solubilité de l'eau dans le méthyl tert-butyl-éther est donnée à 5.5% molaire à 52.04°C. Le domaine liquide-liquide complet n'est pas connu à température ambiante, mais cette valeur donne une idée de la quantité maximale d'eau dans le méthyl tert-butyl-éther, et inversement, que l'on peut retrouver dans la phase organique à pression atmosphérique.

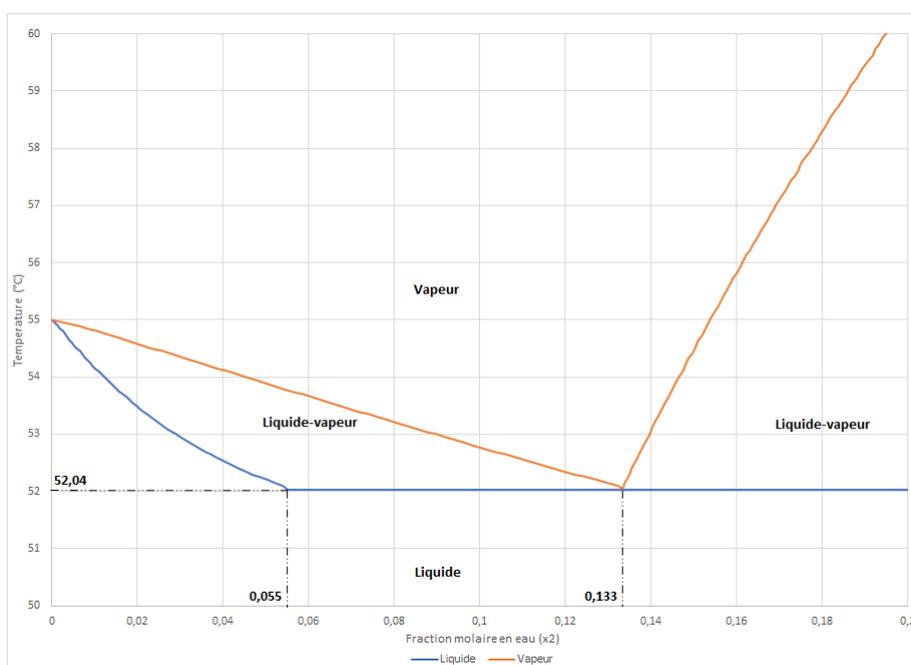


Figure 3.5: Diagramme de phase méthyl tert-butyl-éther/eau à pression atmosphérique

La composition à l'équilibre thermodynamique entre les deux phases est donnée par l'isotherme correspondant. Par exemple, à 52.04°C, une composition liquide à 5.5% molaire en eau donnera une phase vapeur dont la composition correspondra au point azéotrope, soit 13.3% molaire en eau. Autre exemple, à 53°C la composition liquide sera environ de 3% molaire en eau et la phase vapeur 9%. Ainsi au fur et à mesure de la distillation, les fluides voient leur température augmenter jusqu'à atteindre le point d'ébullition du méthyl tert-butyl-éther pur, soit 55°C. Bien évidemment, obtenir du méthyl tert-butyl-éther de très haute pureté est une opération délicate car les compositions des deux phases à l'approche de ce point deviennent équivalentes, donc moins discriminantes vis-à-vis de l'eau. Si l'élimination de l'eau dans la phase liquide est aisée en début de distillation, elle demandera beaucoup plus d'efforts en fin d'opération, soit plus de solvant frais à employer.

Pour la distillation considérée ici, les compositions massiques initiales et finales attendues sont résumées dans le tableau 3.1 en tenant compte du principe actif en solution. La spécification en eau est fixée à 0.3% en masse, soit 3000 ppm en partant d'un mélange à 2.10%. Concernant la teneur finale en principe actif, elle est déterminante car elle conditionne le comportement de la cristallisation en aval du

procédé, deux spécifications sont alors possibles :

- La spécification basse : 30% massique
- La spécification haute : 32% massique

Tableau 3.1: Compositions massiques à l'état initial et final pour la distillation du principe actif

	Eau		MTBE		Principe actif		Total	
	kg	% masse	kg	% masse	kg	% masse	kg	% masse
Initial	1.82	2.10	58.85	67.9	26	30	86.67	100
Spécification basse	0.26	< 0.3	60.40	> 69.7	26	30	86.67	100
Spécification haute	0.24	< 0.3	55.01	> 67.7	26	32	81.25	100

Deux spécifications sont à respecter, la teneur en eau et celle en principe actif. Ce dernier ne se vaporisant pas pour de telles conditions opératoires, ou du moins de façon négligeable, on peut alors considérer uniquement l'eau et le méthyl tert-butyl-éther en supposant l'effet négligeable du soluté sur le système binaire. Le tableau 3.2 résume les pourcentages et les quantités molaires du milieu sans la présence du principe actif. La limite maximale en eau est alors donnée pour 2.07 à 2.13% molaire selon la concentration finale en principe actif obtenue.

Tableau 3.2: Compositions molaires à l'état initial et final pour la distillation du principe actif sans compter la présence du soluté

	Eau		MTBE		Total	
	mol	% mol	mol	% mol	mol	% mol
Initial	101.0	13.14	667.6	86.86	768.6	100
Spécification basse	14.5	2.07	685.3	97.93	699.7	100
Spécification haute	13.6	2.13	624.0	97.87	637.6	100

On étudiera alors deux méthodes de distillation simples, la distillation batch, actuellement employée, et la distillation continue. Le but étant d'éliminer l'eau du mélange en gardant une quantité fixe de méthyl tert-butyl-éther pur. Les résultats issus du modèle seront comparés à ceux générés lors de la campagne pilote à UCB Pharma. Par la suite le modèle continu sera évalué pour une possible implémentation à cette échelle.

3.2.2 Distillation simple batch

Modèle

La distillation batch est la plus ancienne technique de séparation et est très utilisée dans l'industrie pharmaceutique. Son avantage majeur est sa flexibilité qui permet de gérer facilement les stocks de matières et leurs spécifications. Il est admis que la vapeur formée durant une courte période de temps est en équilibre avec sa phase liquide. Ainsi, il existe une relation entre la composition de la phase vapeur y et celle de la phase liquide x qui peut être donnée par la thermodynamique (voir chapitre 2) en fonction de la pression et de la température [65].

L'analyse de cette méthode, présentée par Rayleigh en 1902, marque le début des travaux sur la distillation batch, elle est alors appelée "distillation de Rayleigh" ou "distillation différentielle". Il faut s'intéresser plus particulièrement à l'état initial et final plutôt qu'à la dynamique des flux engagés. L'équation de conservation de la matière pour toute la durée de l'opération est présentée en (3.1) avec S la quantité totale de matière engagée représentant la somme de la quantité dans le réacteur R_{final} avec la quantité de distillats D_{total} [65].

$$S = R_{final} + D_{total} \quad (3.1)$$

Concernant la balance de l'espèce la plus volatile, on réutilise l'équation (3.1) en y ajoutant sa fraction molaire totale x_{total} , sa fraction molaire restée dans le réacteur $x_{R,final}$ et sa fraction molaire moyenne dans les distillats y_{moy} pour donner l'équation (3.2).

$$x_{total} S = x_{R,final} R_{final} + y_{moy} D_{total} \quad (3.2)$$

On a alors trois inconnues pour un système binaire et pour une distillation simple batch : R_{final} , D_{total} et y_{moy} et seulement deux équations, (3.1) et (3.2). Une dernière relation est alors nécessaire pour résoudre ce problème, l'équation de Rayleigh qui considère que :

- la quantité de distillat sortant du réacteur pour un temps court, dD , est égale à la quantité perdue par le liquide dans ce même réacteur, $-dR$ (équation (3.3))
- la quantité en composé le plus volatile sortant du réacteur pour un temps court, $y dD$, est égale à la quantité perdue par le réacteur pour ce même composé, $-d(R x_R)$ (équation (3.4))

$$dD = -dR \quad (3.3)$$

$$y dD = -d(R x_R) \quad (3.4)$$

En combinant ces équations, on obtient la relation (3.6) qui après intégration aboutit à l'équation de Rayleigh (3.7).

$$\begin{aligned} -y dR &= -d(R x_R) \\ -y dR &= -R dx_R - x_R dR \\ R dx_R &= (y - x_R) dR \end{aligned} \quad (3.5)$$

$$\frac{dx_R}{y - x_R} = \frac{dR}{R}$$

$$\int_S^{R_{final}} \frac{dR}{R} = \int_{x_{total}}^{x_{R,final}} \frac{dx_R}{y - x_R} \quad (3.6)$$

$$\ln \left(\frac{R_{final}}{S} \right) = \int_{x_{total}}^{x_{R,final}} \frac{dx_R}{y - x_R} \quad (3.7)$$

Afin de pouvoir intégrer la relation (3.6), la fraction molaire du composé le plus volatile de la phase vapeur, y , doit être une fonction de la fraction de ce même composé dans la phase liquide du réacteur, x_R . Il est donc indispensable de poser l'hypothèse que la phase vapeur est en équilibre thermodynamique avec la phase liquide pour pouvoir employer le diagramme de phases généré précédemment (figure 2.7). De cette manière, pour chaque composition liquide, x , entre la valeur initiale, x_{total} , et finale, $x_{R,final}$, la

valeur de la composition vapeur, y , peut être retrouvée (figure 3.5). Plus concrètement, si l'on se réfère à l'équation de Rayleigh en (3.7), il est possible de construire un graphique avec $\frac{1}{y-x}$ en fonction de x ce qui conduit à la figure 3.8 présentée dans la partie suivante. La fonction $\frac{1}{y-x}$ n'a aucune signification physique.

Enfin, on peut résumer une distillation batch comme présenté sur la figure 3.6. Le système voit son volume diminuer au cours du processus, comme une opération de concentration et où l'un des deux composés du système binaire, ici l'eau, passe préférentiellement dans la phase vapeur que dans la phase liquide. Le deuxième composé, ici le MTBE, est alors concentré dans la phase liquide.

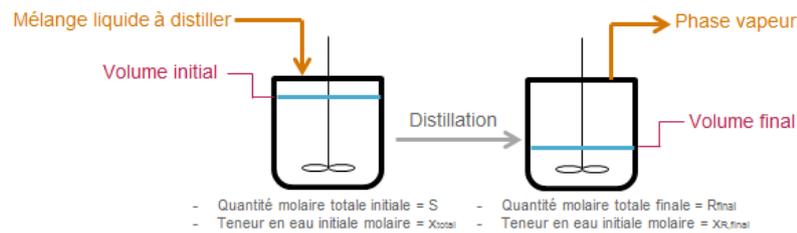


Figure 3.6: Schéma d'un exemple d'une distillation de type batch

Dans la pratique, la quantité molaire totale finale, R_{final} , ou bien le volume final est imposé. En d'autres termes, il faudra adapter judicieusement le volume initial de telle manière qu'une fois le volume final atteint, la quantité résiduelle en solvant à éliminer soit en même temps en dessous des spécifications définies ($< 0.3\%$ en masse), sinon il faudra effectuer une nouvelle distillation.

Résultats

L'ensemble du procédé de la distillation batch est présenté par la figure 3.7, regroupant les conditions initiales et finales. Après ajout de méthyl tert-butyl-éther frais dans le milieu de départ, l'eau est par conséquent diluée, passant de 2.1 à 1.0% en masse, correspondant à 5.5% molaire sans principe actif et cet état sera le point de départ des calculs effectués ci-après.

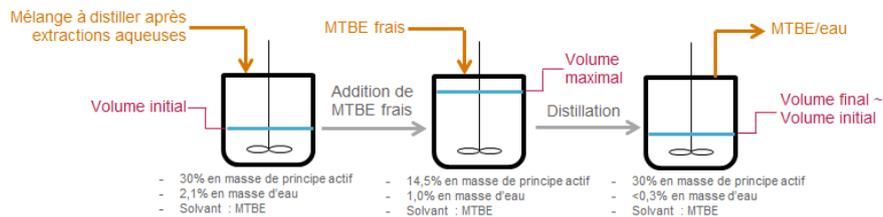


Figure 3.7: Schéma de la distillation de type batch réalisée à échelle pilote à UCB Pharma

Comme présenté dans le tableau 3.2, la quantité molaire totale visée se situe entre 699.7 et 637.6 moles sans prendre en compte la présence du principe actif et en partant d'un mélange de 768.6 moles auquel on ajoute 1051 moles de méthyl tert-butyl-éther frais. La démarche employée consiste d'abord à réconcilier les valeurs des quantités totales molaires et ensuite de comparer les teneurs en eau obtenues

par le modèle. La valeur de l'intégration entre x_{total} et $x_{R,final}$ est utilisée dans l'équation de Rayleigh pour déduire numériquement la quantité de liquide qui restera dans le réacteur, R_{final} (équation (3.9)).

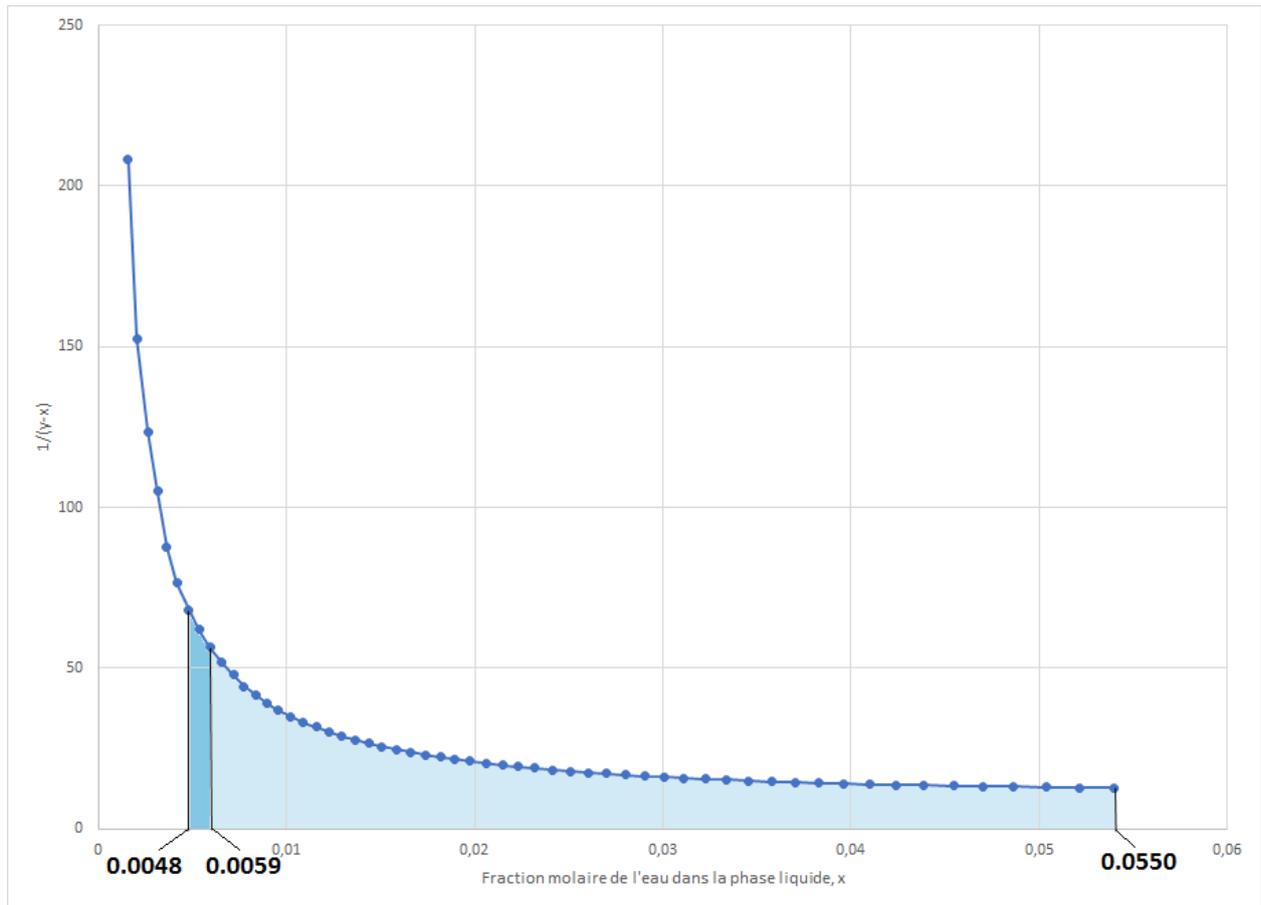


Figure 3.8: Graphique de $\frac{1}{y-x}$ en fonction de la fraction molaire en eau x_R avec l'aire sous la courbe considérée pour la distillation batch

Par essai-erreur, on obtient deux valeurs proches de ces limites spécifiées : 686.0 moles et 638.8 moles correspondant respectivement à des teneurs en eau de 0.59% et 0.48% molaire, résumées dans le tableau 3.3 et par la figure 3.8. On voit facilement que l'aire sous la courbe à l'approche de zéro devient conséquente, ceci s'explique par des valeurs de plus en plus proches des compositions des deux phases, x et y . Les dernières traces d'eau dans le méthyl tert-butyl-éther deviennent alors très difficiles à éliminer, et de façon générale ce constat s'observe aussi pour tous les systèmes de solvants. C'est pourquoi l'impact sur les quantités de méthyl tert-butyl-éther frais à utiliser devient conséquent lorsqu'il s'agit de faibles variations de quantité d'eau et pour des concentrations basses.

Tableau 3.3: Compositions molaires finales théoriques pour la distillation batch du principe actif et sans compter la présence du soluté

	Eau		MTBE		Total		MTBE frais utilisé mol
	mol	% mol	mol	% mol	mol	% mol	
Limite basse	4.0	0.59	682.0	99.41	686.0	100	1051.0
Limite haute	3.1	0.48	635.7	99.52	638.8	100	1051.0

Si on prend par exemple la spécification basse, l'aire sous la courbe entre x_{total} et $x_{R,final}$ est donnée numériquement en 3.8.

$$\int_{0.055}^{0.0059} \frac{dx_R}{y - x_R} = -1.21159 \quad (3.8)$$

Cette valeur intégrée détermine alors la quantité totale restante de liquide, R_{final} , dans le réacteur, sans compter la présence du principe actif (3.9).

$$R_{final} = S \exp\left(\int_{0.055}^{0.0059} \frac{dx_R}{y - x_R}\right) = 686.0 \text{ mol} \quad (3.9)$$

Au final, le modèle prédit une teneur en eau bien inférieure à la spécification imposée qui est de 2.07-2.13% molaire. Si ces valeurs sont confirmées par les résultats à échelle pilote, il serait alors possible de réduire la quantité de méthyl tert-butyl-éther employée pour atteindre cette limite. Dans la section suivante, on étudiera la distillation en continu. En modifiant la technique de distillation, on cherchera à diminuer davantage l'emploi des solvants.

3.2.3 Distillation simple continue

Modèle

La seule différence entre la distillation simple batch et continue est qu'un flux continu de solvant "frais" ou de pureté très élevée est apporté au sein du réacteur afin de maintenir constant le niveau de liquide dans le réacteur, elle est alors aussi appelée "distillation à volume constant" (figure 3.9). Comparativement à la distillation batch, l'avantage majeur ici est de ne pas avoir à adapter le volume initial avant distillation. L'opération peut être interrompue à tout instant car la teneur en principe actif est considérée quasi constante, seule la teneur en eau est impactée. On a ainsi un seul paramètre qui varie, à l'inverse de la distillation batch qui demande de concilier deux paramètres en même temps, ce qui rend ce procédé plus flexible et adaptable. Des échantillons peuvent aussi être pris à tout moment afin de suivre cette concentration et d'arrêter le processus au bon moment.

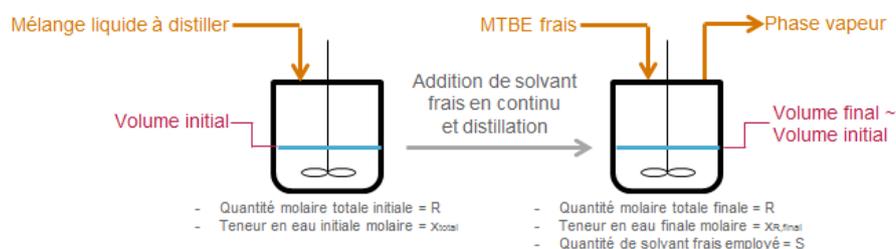


Figure 3.9: Schéma d'une distillation de type continue

Il faut noter que c'est la quantité totale molaire totale dans le réacteur qui est maintenue constante, on pose alors l'approximation que le volume restera constant également. La balance molaire est donnée par

la relation (3.10) qui se simplifie donc aisément.

$$\begin{aligned} [Entree] - [Sortie] &= [Accumulation \text{ dans le reacteur}] \\ &= 0 \end{aligned} \quad (3.10)$$

Pour maintenir le niveau constant, la quantité de solvant évaporé pour un temps court, $-dV$, doit être égale à la quantité de solvant ajouté, dS (équation (3.11)).

$$-dV = dS \quad (3.11)$$

Si l'on reprend les notations de la section 3.2.2, la balance sur le composé le plus volatile est donnée par l'équation (3.12).

$$\begin{aligned} -y dV &= -d(R x_R) \\ y dV &= d(R x_R) \\ y dV &= R dx_R + x_R dR \end{aligned} \quad (3.12)$$

Comme la quantité totale, R , est gardée constante on peut simplifier la relation établie en (3.12), la substituer avec (3.11) et intégrer l'équation obtenue (démonstration (3.13)).

$$\begin{aligned} y dV &= R dx_R \\ -y dS &= R dx_R \\ \frac{dS}{R} &= -\frac{dx_R}{y} \\ \int \frac{dS}{R} &= -\int_{x_{R,initial}}^{x_{R,final}} \frac{dx_R}{y} \\ \frac{S}{R} &= -\int_{x_{R,initial}}^{x_{R,final}} \frac{dx_R}{y} \end{aligned} \quad (3.13)$$

De la même manière que la distillation batch, les deux phases sont supposées à l'équilibre thermodynamique et la résolution graphique est également possible avec $\frac{1}{y}$ en fonction de x_R (figure 3.11). Encore une fois, la fonction $\frac{1}{y}$ n'a pas de sens physique, elle est utilisée pour déduire l'aire sous la courbe à partir de l'équation (3.13).

Cette méthode n'a pour le moment pas été utilisée à grande échelle, la partie suivante s'intéressera alors à reproduire les mêmes résultats que la distillation batch et de comparer les quantités de méthyl tert-butyl-éther employées.

Résultats

Comme développé dans la partie précédente, la distillation continue se fait à volume constant, on étudiera uniquement le cas de la spécification basse, c'est à dire garder un milieu contenant 30% de principe actif en masse puisque le milieu de départ contient cette même concentration (figure 3.10).

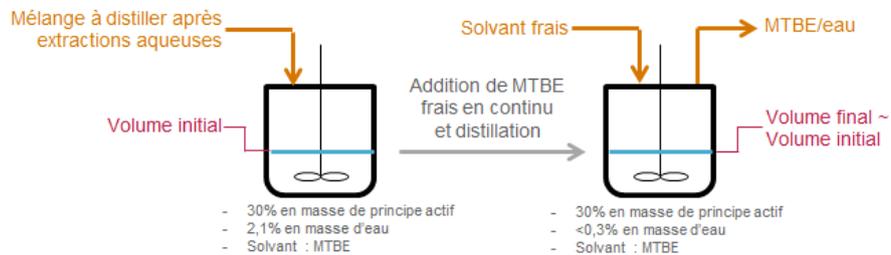


Figure 3.10: Schéma de la distillation de type continu proposée pour l'échelle pilote

Dans le cas de cette distillation, le milieu de départ est donc différent de la distillation batch car non dilué. La composition initiale correspond donc à celle juste après les lavages aqueux, c'est à dire 13.14% molaire en eau (tableau 3.2). On construit alors la courbe 3.11 entre]0; 13.14] afin de représenter l'aire qui sera utile au calcul décrit dans la partie précédente.

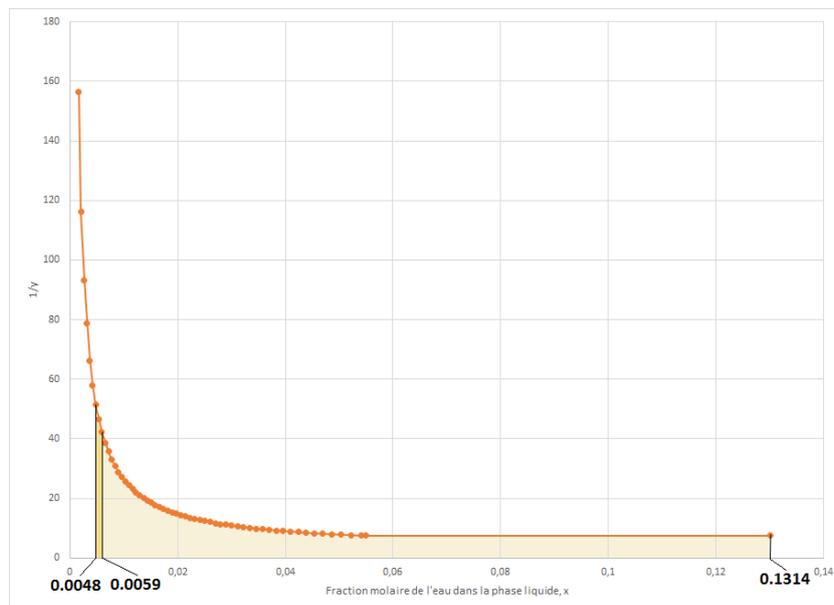


Figure 3.11: Graphique de $\frac{1}{y}$ en fonction de la fraction molaire en eau x_R avec l'aire sous la courbe considérée pour la distillation continue

L'aire contenue entre 13.14 et 0.59% molaire est calculée numériquement et introduite dans l'équation de Rayleigh déterminée en 3.13 via l'expression 3.14.

$$\left(\int_{0.1314}^{0.0059} \frac{dx_R}{y} \right) = -1.2347 \quad (3.14)$$

$$S = R \left(\int_{0.1314}^{0.0059} \frac{dx_R}{y} \right) = 863.9 \text{ mol} \quad (3.15)$$

Les résultats sont résumés dans le tableau 3.4 pour la spécification basse. La quantité de méthyl tert-butyl-éther S nécessaire à l'obtention d'un milieu à 0.59% molaire en eau est de 863.9 moles (équation (3.15)). Pour rappel, la distillation batch utilisait 1051 moles de solvant frais pour ce même résultat, soit une réduction de 17.8% avec la distillation continue. Dans les deux cas, c'est seulement 96.9 moles d'eau qui sont distillées de la phase liquide. Il faut donc 9 molécules de MTBE pour éliminer 1 molécule d'eau pour la distillation continue contre 11 molécules de MTBE pour la distillation batch.

Tableau 3.4: Compositions molaires finales théoriques pour la distillation continue du principe actif et sans compter la présence du soluté

Eau		MTBE		Total		MTBE frais utilisé mol
mol	% mol	mol	% mol	mol	% mol	
4.1	0.59	695.6	99.41	699.7	100	863.9

La figure 3.12 montre le profil de la concentration en eau en fonction de la quantité de distillats récoltés pour les deux méthodes. On voit clairement une meilleure efficacité de la méthode continue par rapport à celle discontinue pour l'élimination de l'eau résiduelle.

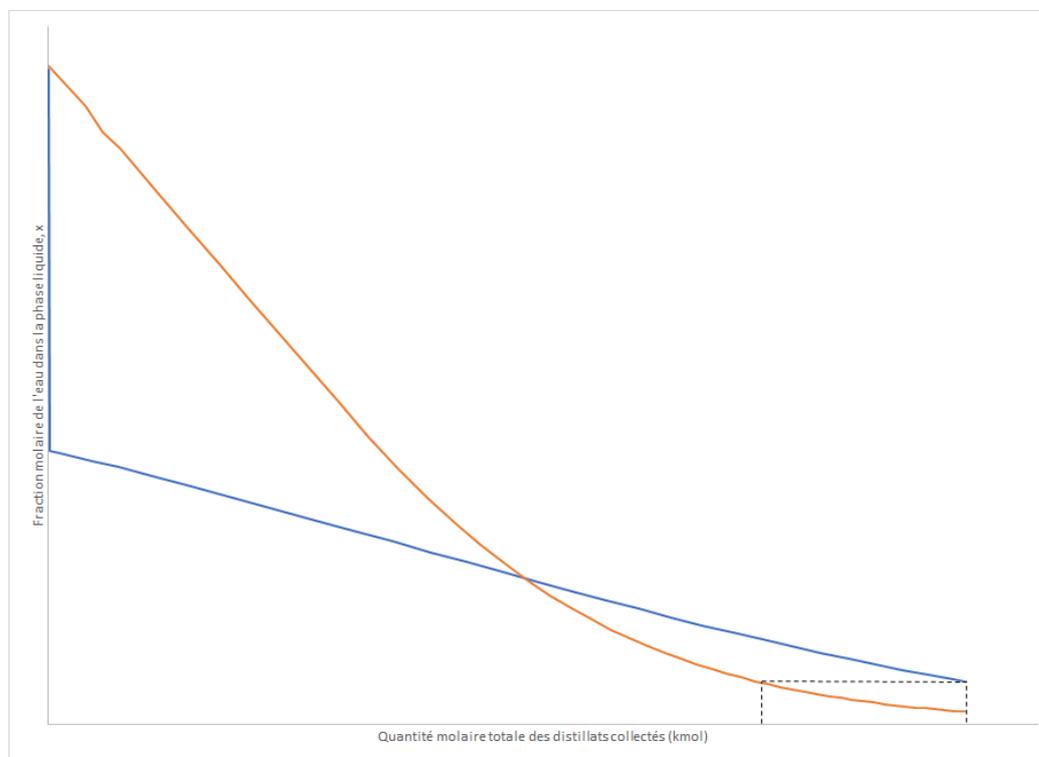


Figure 3.12: Graphique comparatif de la fraction molaire en eau dans la phase liquide en fonction de la quantité molaire de distillats collectés pour la distillation continue (en orange) et la distillation batch (en bleu) selon les estimations théoriques

De façon générale, les procédés continus peuvent être comparés à une cascade de procédés discontinus. Une faible fraction de solvant est ajoutée dans le réacteur, puis distillée et ainsi de suite. Ainsi, il serait possible avec plusieurs distillations batch à la suite de tendre vers le résultat de la distillation continue, qui est le cas idéal. Pour des raisons techniques et pratiques, cette solution peut être employée si la distillation continue n'était pas réalisable à l'échelle. La quantité totale de solvant serait donc fractionnée et plusieurs distillations batch seraient réalisées à la suite. Par la théorie, on montre ici que l'introduction de l'entièreté du méthyl tert-butyl-éther dans le milieu comme effectué dans la section 3.2.2 représente le cas le moins efficace pour l'élimination de l'eau.

Il faut donc maintenant confronter ces analyses théoriques avec la pratique. Les résultats à l'échelle pilote permettront de valider ou non les résultats du modèle développé jusqu'ici et ceci uniquement pour la distillation batch.

3.2.4 Résultats de la campagne à échelle pilote

Une campagne de production a été effectuée en 2017 pour produire environ deux lots de 25 kilogrammes de principe actif chacun afin d'approvisionner des études cliniques. La distillation continue n'était pas encore envisagée à ce stade, le procédé impliquait une distillation batch à l'origine. Aucun échantillon n'est pris avant la distillation, la composition initiale n'est donc pas connue mais a été évaluée dans les sections précédentes. La composition finale est en revanche connue grâce à l'échantillonnage et doit respecter les spécifications développées dans le tableau 3.2 précédent. Les deux paramètres critiques sont donc la concentration en principe actif et en eau, l'ensemble des résultats pour ces deux lots est présenté dans le tableau 3.5.

Tableau 3.5: Compositions molaires finales, sans tenir compte de la présence du soluté, après distillation pour deux batchs de principe actif réalisés à l'échelle pilote à UCB Pharma

	Eau		MTBE		Total		MTBE frais utilisé mol
	mol	% mol	mol	% mol	mol	% mol	
Batch 1	8.0	1.17	676.1	98.83	684.1	100	1049.3
Batch 2	8.0	1.30	631.3	98.74	639.3	100	1021.0

Dans chaque cas, la concentration en principe actif a été respectée (entre 30 et 32% en masse). Les teneurs en eau sont nettement plus élevées par rapport aux prédictions du modèle, 1.17 et 1.30% molaire contre 0.48 à 0.59% molaire attendus. Ces résultats sont tout de même largement inférieurs aux limites maximales définies qui se situent entre 2.07 et 2.13% molaire présentés par le tableau 3.2. Quatre explications sont proposées pour expliquer ces écarts, elles sont détaillées ci-après.

Quantité réelle de solvant frais utilisée

La quantité de solvant frais introduite avant distillation est inférieure à celle initialement prévue, 1051 moles de méthyl tert-butyl-éther étaient prévues dans le modèle, ici seulement 1049.3 et 1021.0 moles ont été utilisées. Cependant ces différences ne sont pas assez significatives pour justifier un tel écart dans la concentration en eau. On a vu précédemment qu'une faible variation de la concentration en eau implique une grande quantité de solvant surtout pour des concentrations très basses. Or, ici, 30 moles de

différence ne représentent que 3.6 litres de solvant par rapport à 122 litres introduits. Cette explication n'est donc pas suffisante pour expliquer ces résultats.

Phénomène de reflux

Le phénomène de reflux dans le réacteur est un aléa de la distillation simple. L'hypothèse de départ est de considérer que tout le réacteur est à température de distillation et de façon homogène. Dans la pratique, seul le fond et la circonférence du réacteur est couvert d'une double enveloppe dans laquelle le liquide caloporteur apporte l'énergie au système (voir figure 3.13).

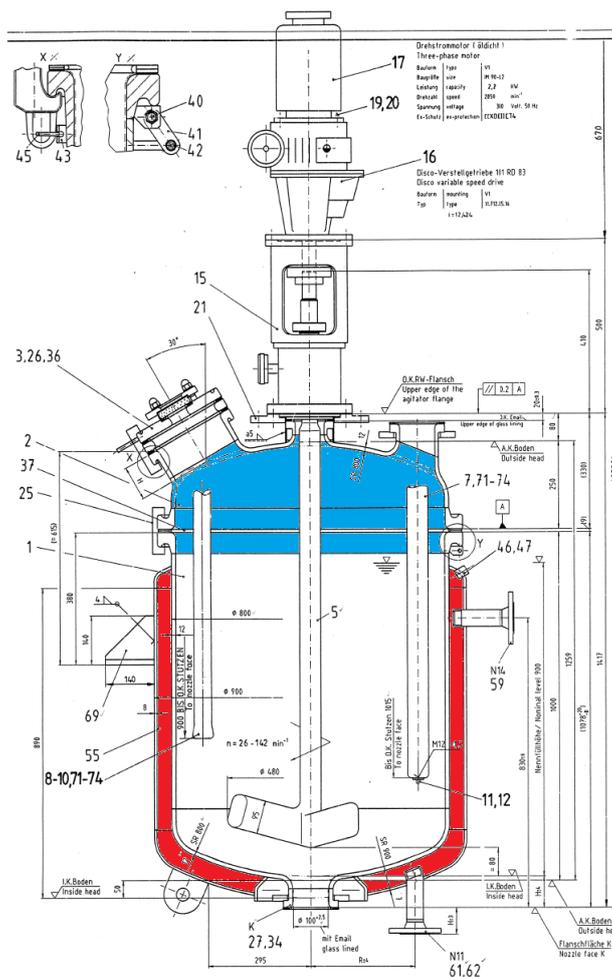


Figure 3.13: Dessin technique du réacteur Pfaudler AE400 utilisé pour la distillation, en rouge la partie couverte par la double enveloppe, en bleu la partie "froide" du réacteur

Les vapeurs issues du liquide distillé s'échappent par le haut du réacteur et rencontrent le dôme qui a une température plus basse que le reste du réacteur. Il en résulte une condensation partielle des vapeurs sur les parois et en faveur du composé qui a la volatilité la plus haute, ici l'eau. La fraction gazeuse de méthyl tert-butyl-éther est donc plus importante au détriment de l'eau, ce qui n'est pas souhaité ici pour ce procédé.

A l'inverse dans le cas d'une distillation fractionnée c'est un phénomène très exploité, la colonne de séparation permettant de séparer les composés du mélange selon leur volatilité, notamment pour la séparation des hydrocarbures du pétrole brut. Pour la distillation effectuée dans le réacteur AE400, il est difficile de quantifier le reflux car il est inhérent au design du réacteur. Au final, il est tout à fait possible d'affiner les résultats en introduisant un tel paramètre dans le modèle, mais sa part de responsabilité pourrait se révéler moins importante que prévue.

Exactitude du modèle NRTL

Le modèle NRTL est un modèle quasi-empirique et non exact du point de vue mécanistique. Tous les calculs effectués jusqu'ici dépendent initialement du modèle thermodynamique qui a généré le diagramme des équilibres liquide-vapeur. La partie du diagramme qui est exploitée est aussi très restreinte si l'on se réfère à la figure 2.7. Ainsi, une faible erreur de la part du bioalgorithme lors de la minimisation de l'énergie libre de Gibbs pour générer les compositions liquides et vapeur, x et y , suffit à entraîner des erreurs numériques conséquentes dans l'intégration de Rayleigh.

Effet du principe actif sur la distillation

Nous avons considéré que le principe actif n'avait aucun impact sur la distillation. Cette hypothèse n'est pas tout à fait exacte dans la pratique. En effet, on sait déjà que ce produit solubilise plus d'eau dans la phase organique qu'estimé, 13.14% contre 5.50% molaire estimé selon le diagramme. De plus, cette concentration initiale en eau n'est pas connue à cette échelle et a été uniquement déterminée au laboratoire. On peut imaginer également que le point azéotrope est modifié par cette présence, ce qui remet en cause la pertinence du diagramme. L'analyse en ligne des compositions du liquide et des distillats dans un réacteur parfaitement calorifugé au laboratoire, afin de s'affranchir de l'effet de reflux présenté précédemment, permettrait d'évaluer cet écart.

3.3 Pistes d'amélioration

3.3.1 Amélioration du modèle

Le modèle présenté a permis de donner une première estimation des résultats à l'échelle pilote. Beaucoup d'hypothèses simplificatrices ont été posées pour son élaboration, notamment le fait de négliger l'effet du principe actif sur les équilibres thermodynamiques pour le binaire eau/méthyl tert-butyl-éther. Dans ce cas précis, il serait utile de générer de nouveaux paramètres NRTL :

- entre l'eau et le principe actif
- entre le méthyl tert-butyl-éther et le principe actif

Avec trois nouveaux ensembles de paramètres il serait alors possible de construire un diagramme ternaire des équilibres liquide-vapeur, dont la forme générale est présentée par la figure 3.14.

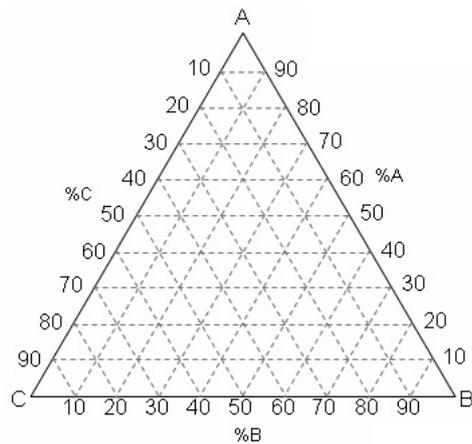


Figure 3.14: Diagramme ternaire vierge des équilibres liquide-vapeur pour trois composés A, B et C

Il faudrait cependant effectuer un travail considérable au laboratoire pour fournir suffisamment de données expérimentales utiles à la régression de ces paramètres. Le bioalgorithme présenté dans la première partie devrait également être modifié en conséquence pour construire des diagrammes ternaires et requiert un temps de calcul bien plus long que pour un diagramme binaire de par sa dimension supplémentaire.

Comme évoqué dans la deuxième partie, il serait également intéressant d'employer le modèle UNIFAC ou UNIQUAC qui se base sur les contributions de groupes chimiques. Ainsi, le comportement du principe actif pourrait être estimé vis-à-vis de l'eau et du méthyl tert-butyl-éther. Il est possible aussi de régresser les données issues du modèle UNIFAC ou UNIQUAC en paramètres NRTL. Moins de travail expérimental serait bien évidemment requis mais complexifierait le modèle thermodynamique. Souvent, de tels modèles sont souvent compromis à cause des comportements erratiques des molécules pharmaceutiques, celles-ci ayant des structures complexes qui rendent la modélisation compliquée. L'impact mérite d'être évalué.

3.3.2 Amélioration du procédé

A travers cette modélisation, il a pu être mis en évidence que la distillation simple batch employée dans ce procédé n'est pas adaptée et peu efficace pour l'élimination de l'eau solubilisée dans la phase organique. Alternativement, la distillation continue proposée montre des résultats théoriques bien plus intéressants par comparaison avec la distillation actuelle. Elle a l'avantage d'être plus simple dans sa maîtrise car elle peut être interrompue à tout moment et demande moins de volume opératoire. De plus, la consommation de solvant serait nettement réduite et contribuerait à une réduction des coûts et de l'impact écologique. Cette nouvelle procédure a été proposée au sein d'UCB Pharma afin d'être implémentée dans la prochaine campagne qui aura lieu en 2019.

Afin d'accroître la performance de la séparation des deux composés, un autre paramètre opératoire a été aussi considéré : la pression. En effet, elle a un rôle majeur dans la distillation et peut être utilisée pour favoriser l'élimination de l'eau dans le méthyl tert-butyl-éther.

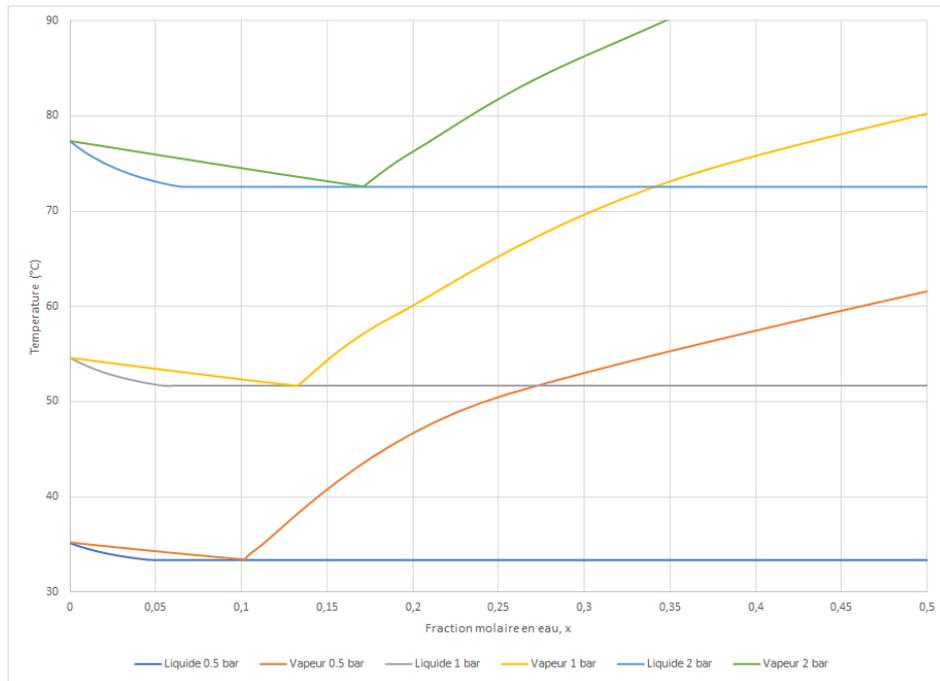


Figure 3.15: Diagramme binaire des équilibres liquide-vapeur pour l'eau et le méthyl tert-butyl-éther à 0.5, 1.0 et 2.0 bar générés par l'algorithme Krill Herd

La figure 3.15 présente trois diagrammes à différentes pressions, on voit facilement qu'en dessous de la pression atmosphérique, la composition azéotrope se rapproche de celle du méthyl tert-butyl-éther pur et rend la séparation de l'eau compliquée, l'écart entre les compositions liquide-vapeur étant de plus en plus mince dans la partie du diagramme considérée. A l'inverse, il semble que pour des pressions au-delà de la pression atmosphérique, la tendance est favorable à une bonne élimination de l'eau, la composition azéotrope se décalant vers la droite. L'économie de solvant frais serait encore plus importante si la pression était augmentée au sein du réacteur. Malheureusement, pour des raisons techniques, il n'est pas possible d'appliquer une pression supérieure à celle atmosphérique de quelques millibars dans ce réacteur. Cette solution n'a donc pas été retenue pour la prochaine campagne.

Conclusion

La modélisation des procédés de fabrication au sein de l'industrie pharmaceutique est une pratique non exploitée pleinement mais en plein essor car elle dispose d'un bon potentiel prédictif tout en restant intéressant économiquement pour l'entreprise. Beaucoup plus employée dans la pharmacie galénique, elle peine encore à se faire une place dans les procédés de synthèse. De plus, les molécules issues de synthèses non biologiques deviennent plus compliquées à produire, les synthèses étant nombreuses et les rendements pas toujours satisfaisants. En conséquence, de telles molécules deviennent de plus en plus coûteuses lors de la phase de développement pour une entreprise pharmaceutique, en particulier lorsque les méthodes de travail restent empiriques. La modélisation est donc un atout à jouer dès la phase de développement, car basée sur des connaissances multi-disciplinaires et bénéficiant de données accumulées depuis des décennies, elle apporte une compréhension profonde des procédés de fabrication afin de les rendre adaptés, flexibles et robustes, tout en limitant le travail expérimental qui peut s'avérer être long. Sur cette base, le temps employé à l'exploration des paramètres critiques est réduit et les problématiques à grande échelle sont résolues plus rapidement. Le début de vie d'un médicament est une phase où l'apport de modifications sur le procédé se révèle moins contraignant d'un point de vue réglementaire comparativement à une molécule qui voit sa commercialisation approcher : cet avantage est à exploiter pour mettre en oeuvre les meilleures solutions à l'étude.

Ce travail décrit les étapes clés permettant la modélisation d'une opération unitaire en vue de l'améliorer et de justifier les modifications proposées à l'aide d'arguments théoriques. Un outil de résolution mathématique relativement simple a été entièrement conçu en vue de résoudre des problèmes de thermodynamique associés aux phénomènes rencontrés lors de cette opération. L'évaluation de la performance et de la robustesse de ce bioalgorithme s'est révélée satisfaisante pour l'appliquer à ce type de problématique. Beaucoup d'outils similaires et plus complexes existent également et peuvent être acquis dans le commerce pour des coûts non négligeables, la solution proposée ici montre qu'il est possible de s'en affranchir en contrepartie d'un effort initial de conception, ce qui est avantageux pour l'entreprise à long terme.

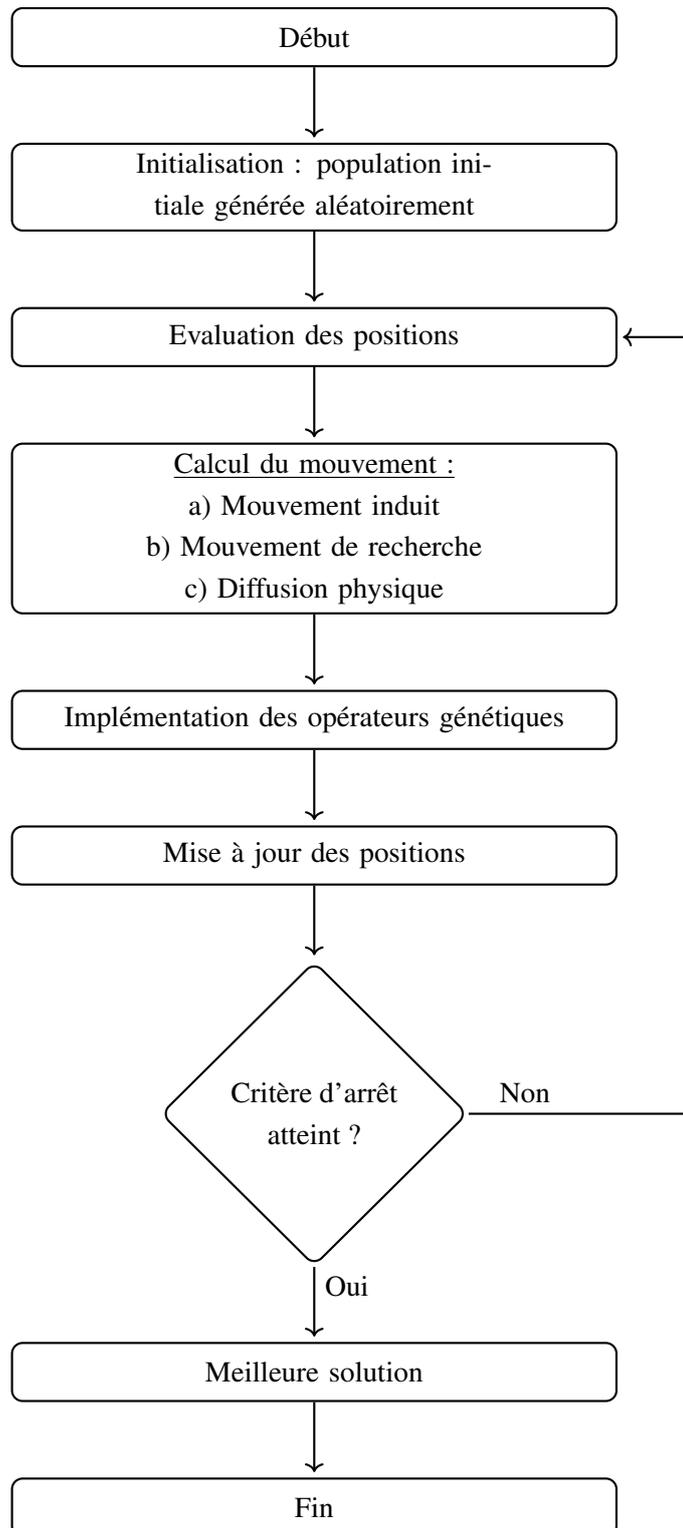
Pour le système considéré dans ce document, des modèles et des hypothèses simplificatrices sont choisis puis traités et évalués. D'abord, les résultats issus de la production à grande échelle sont confrontés avec ceux du modèle pour juger de leur pertinence. L'analyse de ceux-ci ont conduit alors à apporter des critiques sur le choix des hypothèses et d'évaluer les paramètres à inclure par la suite. La compréhension des mécanismes qui influent sur les résultats est un premier travail essentiel et peut être relié à divers paramètres, comme le design spécifique des réacteurs, le comportement des substances ou bien à la réalité d'exécution sur le terrain. Cette remise en question aboutit à une amélioration continue des modèles tout en accumulant de l'expérience sur le procédé avant la mise sur le marché du produit.

Ensuite, même si un modèle ne peut donner qu'une première estimation peu exacte, il est possible à travers cette démarche d'orienter rapidement la recherche vers la voie la plus adaptée selon les contraintes imposées. Dans l'exemple traité ici, il est ainsi intéressant de pouvoir démontrer qu'une distillation continue sera dans tous les cas plus avantageuse à suivre qu'une distillation batch puisqu'elle permettrait de :

- diminuer la consommation de solvant et donc de diminuer l'impact environnemental induit par les déchets générés
- simplifier l'opération en manipulant moins de variables
- diminuer le volume maximal du procédé qui au final conditionne les tailles de réacteurs à utiliser

La mise en place de telles démarches anticipées dans le développement d'un médicament permettent de systématiser l'amélioration continue et d'aboutir à des procédés avec une certaine maturité avant la commercialisation. La place de l'ingénieur est donc cruciale lorsqu'il s'agit de traiter ces problématiques techniques tout cela avec un certain esprit d'innovation dans l'intérêt de l'environnement et de l'économie de l'entreprise. Etant donné le domaine d'application, les compétences du pharmacien industriel aident à contextualiser les événements afin de discuter de leurs impacts éventuels sur la réglementation ou la qualité du produit. Ainsi il est indispensable de disposer d'une équipe multi-disciplinaire qui implique non seulement ingénieurs et pharmaciens, mais également des chimistes, des analystes ou encore des statisticiens pour améliorer la phase de développement d'un médicament à l'aide d'approches innovantes.

Annexe 1 : Méthodologie appliquée à l'algorithme Krill Herd



Annexe 2 : Algorithme Krill Herd en Fortran 90

```

1  PROGRAM Krill_herd_algorithm
2  ! Ce programme a pour but de minimiser des fonctions à l'aide de l'algorithme Krill Herd.
3  ! Le Krill Herd Algorithm est un algorithme basé sur la simulation du comportement biologique des krills.
4  ! L'objectif d'un banc de krill est d'augmenter sa densité (minimiser la distance entre les krills) et de
se diriger au plus proche d'un point de nourriture (ici le point minimal)
5  ! Le mouvement d'un krill est donc induit par 3 facteurs :
6  ! 1) la position des autres krills
7  ! 2) la recherche de nourriture
8  ! 3) une diffusion aléatoire
9  ! Des opérateurs génétiques sont également ajoutés afin d'optimiser l'algorithme.
10 ! L'objectif est de minimiser l'énergie de Gibbs pour un mélange eau+méthyl-tert-butyl-éther à
température, pression et proportion molaire donnée et de tracer le diagramme de phases correspondant.
11  IMPLICIT NONE
12  !Variables du krill
13  INTEGER::It,Itmax,dim,n,i,j,choix,Cy,Cymax,etat ! Déclaration des variables entières
14  REAL(8)::dt,K,Kfinal,RanNb,Kfood,initmin,initmax,initmean,valmax,valmin,sommean,valmean ! Déclaration
des scalaires/variables doubles
15  REAL::time,Tarray(2) !Variables utilisée pour le temps de calcul
16  REAL(8),ALLOCATABLE::LB(:),UB(:),Xfinal(:),Kvect(:),Kibest(:),Xfood(:,:),Xfin(:,:),Kibestchap(:) !
Déclaration des vecteurs doubles
17  REAL(8),ALLOCATABLE::X(:,:),dX(:,:),N1(:,:),D(:,:),F(:,:),Xnew(:,:),Kmat(:,:),Xibest(:,:),Xmat(:,:,:)!
Déclaration des tableaux doubles
18  !Variables partie thermodynamique
19  REAL(8)::T,z1,P
20  INTEGER::side,sidemax
21  REAL(8),PARAMETER::epsth=1.d-10
22  INTEGER,PARAMETER::methode=1
23
24  !===== RANDOM SERIES=====
25  INTEGER:: nbr, temps(8), heure, minute, seconde, centieme
26  INTEGER,ALLOCATABLE:: seed(:)
27  ! on détermine n, le nombre d'entiers utilisés par défaut
28  ! par le processeur pour générer le seed
29  CALL RANDOM_SEED(SIZE=nbr)
30  ! n étant connu, on dimensionne le tableau seed(:)
31  ALLOCATE(seed(nbr))
32  !lecture de l'horloge interne
33  CALL DATE_AND_TIME(VALUE= temps)
34  heure = temps(5)
35  minute = temps(6)
36  seconde = temps(7)
37  centieme = temps(8)
38  !initialisation de la séquence aléatoire
39  seed = (heure+1)*(minute+1)*(seconde+1)*(centieme+1)
40  ! génération du premier terme de la suite à partir de la
41  ! valeur de seed(:)
42  CALL RANDOM_SEED(PUT=seed)
43
44  !===== END RANDOM SERIES =====
45
46  !===== CONDITIONS INITIALES THERMO
47
48  OPEN(22,FILE='Parametres_thermo.txt') ! Lecture des paramètres réglables dans le fichier texte
'Parametres_thermo.txt'
49  READ(22,*)
50  READ(22,*)
51  READ(22,*)
52  READ(22,*)T ! Récupère une valeur fixe de la température pour le choix 6 (voir plus bas): calcul
thermo en un point
53  READ(22,*)
54  READ(22,*)
55  READ(22,*)z1 ! De même, récupère une valeur fixe de z1 pour le choix 6
56  READ(22,*)
57  READ(22,*)
58  READ(22,*)P ! La valeur de la pression est valable pour toute la thermo (choix 6 et 7)
59  CLOSE(22)

```

```

60
61 !===== INTERFACE UTILISATEUR =====
62     WRITE(*,*)
63     WRITE(*,*) "======"
64     WRITE(*,*) "====KRILL===="
65     WRITE(*,*) "====HERD===="
66     WRITE(*,*) "====ALGORITHM===="
67     WRITE(*,*) "===="
68     WRITE(*,*) "                v4.3 "
69     WRITE(*,*)
70     WRITE(*,*) "Choisissez une fonction a minimiser :"
71     WRITE(*,*)
72     WRITE(*,*) "Low-dimensional functions"
73     WRITE(*,*) "-----"
74     WRITE(*,*) "Choix 1 : Fonction de Goldstein-Price"
75     WRITE(*,*) "Choix 2 : Fonction de Camel Back - 6 Hump"
76     WRITE(*,*)
77     WRITE(*,*) "High-dimensional functions"
78     WRITE(*,*) "-----"
79     WRITE(*,*) "Choix 3 : Fonction d'Ackley"
80     WRITE(*,*) "Choix 4 : Fonction de Rastrigin"
81     WRITE(*,*) "Choix 5 : Fonction Quartic"
82     WRITE(*,*)
83     WRITE(*,*) "Thermodynamique"
84     WRITE(*,*) "-----"
85     WRITE(*,*) "Choix 6 : Minimiser en un point"
86     WRITE(*,*) "Choix 7 : Diagramme de phases"
87     WRITE(*,*)
88
89     DO ! La boucle sert à relancer la question ci dessous lorsque l'utilisateur n'entre pas un chiffre
entre 1 et 7
90         WRITE(*,*)"Quel est votre choix ?"
91         READ(*,*) choix ! Lecture du choix de l'utilisateur selon les propositions ci dessus
92         WRITE(*,*)
93
94     ! Selon le choix de l'utilisateur, le programme peut être amené a minimiser des fonctions avec des
dimensions différentes entres elles,
95     ! des bornes différentes ...etc. Chaque fonction demande une configuration qui lui est propre, on retrouve
les paramètres suivants :
96     ! - sidemax : sert uniquement dans la création du diagramme de phases par le KH (choix 7). C'est la
valeur maximale que peut prendre la variable 'side'.
97     !         Dans le cas du diagramme cette valeur vaut 2 : lorsque 'side' vaut 1 une première partie
des données sont récupérées (à gauche), puis lorsque 'side' vaut 2,
98     !         ce sont les valeurs à droite qui sont récupérées. Hormis le diagramme, cette variable ne
sert pas, on lui donne alors par défaut la valeur 1 pour éviter que la boucle ne se lance plus d'une fois
99     ! - dim : représente la dimension (nombre de variables)
100    ! - LB(j) et UB(j) : ce sont les bornes haute et basse des variables (1,2...j...dim) de la fonction
objectif
101
102    IF(choix==1) THEN ! Choix 1 : Fonction de Goldstein-Price
103        sidemax=1
104        IF(methode==1) THEN
105            dim = 2
106            ALLOCATE(LB(dim))
107            ALLOCATE(UB(dim))
108
109            DO j=1,dim
110                LB(j) = -2.d0 !Borne basse
111                UB(j) = 2.d0 !Borne haute
112            ENDDO
113
114            WRITE(*,*) "Please wait..."
115            WRITE(*,*)
116            WRITE(*,*) "Objectif : f=3.000"
117            EXIT !On sort de la boucle une fois le choix effectué
118

```

```

119         ELSE
120
121     ENDIF
122
123
124     ELSEIF(choix==2) THEN ! Choix 2 : Fonction de Camel Back - 6 Hump
125         sidemax=1
126
127         IF(methode==1) THEN
128             dim = 2
129             ALLOCATE(LB(dim))
130             ALLOCATE(UB(dim))
131
132             DO j=1,dim
133                 LB(j) = -5.d0 !Borne basse
134                 UB(j) = 5.d0 !Borne haute
135             ENDDO
136
137             WRITE(*,*) "Please wait..."
138             WRITE(*,*)
139             WRITE(*,*) "Objectif : f=-1.0316285"
140             EXIT
141
142         ELSE
143
144     ENDIF
145
146     ELSEIF(choix==3) THEN ! Choix 3 : Fonction de Ackley
147         sidemax=1
148
149         IF(methode==1) THEN
150             dim = 20
151             ALLOCATE(LB(dim))
152             ALLOCATE(UB(dim))
153
154             DO j=1,dim
155                 LB(j) = -32.d0 !Borne basse
156                 UB(j) = 32.d0 !Borne haute
157             ENDDO
158
159             WRITE(*,*) "Please wait..."
160             WRITE(*,*)
161             WRITE(*,*) "Objectif : f=0"
162             EXIT
163
164         ELSE
165
166     ENDIF
167
168     ELSEIF(choix==4) THEN ! Choix 4 : Fonction de Rastrigin
169         sidemax=1
170
171         IF(methode==1) THEN
172             dim = 20
173             ALLOCATE(LB(dim))
174             ALLOCATE(UB(dim))
175
176             DO j=1,dim
177                 LB(j) = -5.12d0 !Borne basse
178                 UB(j) = 5.12d0 !Borne haute
179             ENDDO
180
181             WRITE(*,*) "Please wait..."
182             WRITE(*,*)
183             WRITE(*,*) "Objectif : f=0"
184             EXIT

```

```

185
186         ELSE
187
188     ENDIF
189
190 ELSEIF(choix==5) THEN ! Choix 5 : Fonction Quartic
191     sidemax=1
192
193     IF(methode==1) THEN
194         dim = 20
195         ALLOCATE(LB(dim))
196         ALLOCATE(UB(dim))
197
198         DO j=1,dim
199             LB(j) = -1.28d0 !Borne basse
200             UB(j) = 1.28d0 !Borne haute
201         ENDDO
202
203         WRITE(*,*) "Please wait..."
204         WRITE(*,*)
205         WRITE(*,*) "Objectif : f=0"
206         EXIT
207
208     ELSE
209
210     ENDIF
211
212 ELSEIF(choix==6) THEN ! Choix 6 : Minimiser en un point (Thermo)
213     sidemax=1
214
215     IF(methode==1) THEN
216         dim = 2
217         ALLOCATE(LB(dim))
218         ALLOCATE(UB(dim))
219
220         LB(1) = epsth !Borne basse de x11
221         UB(1) = z1-epsth !Borne haute de x11
222         LB(2) = z1+epsth !Borne basse de x12
223         UB(2) = 1.d0-epsth !Borne haute de x12
224
225         WRITE(*,*) "Please wait..."
226         WRITE(*,*)
227         EXIT
228
229     IF(DABS(DABS(theta(1))-DABS(theta(2)))<1.d-1) THEN
230         IF(etat==2) THEN
231             WRITE(*,*) "=====> Monophasique vapeur"
232             WRITE(*,*)
233         ELSE
234             WRITE(*,*) "=====> Monophasique liquide"
235             WRITE(*,*)
236         ENDIF
237     ELSE
238         WRITE(*,*) "=====> Diphasique"
239         WRITE(*,*)
240     ENDIF
241
242     WRITE(*,*)
243     WRITE(*,*) "Temps de calcul : ",time, "s"
244     WRITE(*,*)
245     WRITE(*,*) "'Entree' pour quitter"
246     READ(*,*)
247     STOP
248
249 ELSE
250

```

```

251
252         ENDIF
253
254         ELSEIF(choix==7) THEN ! Choix 7 : Diagramme de phases (Thermo)
255             sidemax=2
256
257             IF(methode==1) THEN
258                 dim = 2
259                 ALLOCATE(LB(dim))
260                 ALLOCATE(UB(dim))
261
262                 WRITE(*,*)"Please wait..."
263                 WRITE(*,*)
264                 EXIT
265
266             ELSE
267
268             ENDIF
269
270         ELSE
271
272         ENDIF
273
274     ENDDO
275
276     WRITE(*,*)
277     OPEN(20,FILE='valeurs_thermo_KH.txt') ! Crée/ouvre un fichier qui va stocker les valeurs utiles à la
construction du diagramme de phases
278     OPEN(30,FILE='iterations_KH.txt')
279     ALLOCATE(Xfin(1,dim))
280
281     !===== KRILL HERD ALGORITHM =====
282
283     ! 'i' représente le krill considéré
284     ! 'j' est une variable : (x1 ,x2,..xj,..xdim)
285     ! 'dim' représente le nombre de variables dans la fonction objectif
286     ! Il est alors possible d'établir une matrice X(n,dim)
287     ! X(i,j) correspondant à une valeur de l'une des variables j pour krill i donné
288
289     OPEN(21,FILE='Parametres_krill.txt') ! Lecture des paramètres réglables dans le fichier texte
'Parametres_krill.txt'
290     READ(21,*)
291     READ(21,*)
292     READ(21,*)
293     READ(21,*)n !Nombre de krills
294     READ(21,*)
295     READ(21,*)
296     READ(21,*)Itmax ! Nombre d'itérations à effectuer
297     READ(21,*)
298     READ(21,*)
299     READ(21,*)Cymax ! Nombre de fois qu'il faudra lancer l'algorithme
300     CLOSE(21)
301
302     DO side=1,sidemax ! Pour diagramme : 'side' = 1 : côté gauche, 'side' = 2 : côté droit. Si hors diagramme,
'sidemax' = 1
303
304     IF(side==1 .AND. choix==7) THEN ! Initialisations pour le côté GAUCHE du diagramme de phases
305
306         Xfin(1,1) = 1.d-3 ! x11
307         Xfin(1,2) = 1.d-2 ! x12
308         z1=2.d-3
309
310         ! Equation d'Antoine pour récupérer la température d'ébullition de l'eau : température de départ
311         ! On se place un peu en dessous du point d'ébullition pour bien être dans l'ELV
312         T = 1722.09d0/(5.1792d0-DLOG10(P))-232.8847d0 + 273.15d0 - 0.05d0
313

```

```

314         WRITE(20,*) 0.d0,0.d0,T
315
316     ELSEIF(side==2 .AND. choix==7) THEN ! Initialisations pour le côté DROIT du diagramme de phases
317
318         Xfin(1,1) = 1.d-3! x11
319         Xfin(1,2) = 2.d-2 ! x12
320         z1=5.d-3
321
322         ! Equation d'Antoine pour récupérer la température d'ébullition du méthyl-tert-butyl-éther :
température de départ
323         ! On se place un peu en dessous du point d'ébullition pour bien être dans l'ELV
324         T = 1097.67d0/(3.9698d0-DLOG10(P))-221.8468d0 + 273.15d0 - 0.05d0
325         WRITE(20,*) 1.d0,1.d0,T
326
327     ELSE ! Initialisations pour le choix 6
328         Xfin(1,1) = z1-1.d-2 ! x11
329         Xfin(1,2) = z1+2.d-2 ! x12
330
331
332     ENDIF
333
334     DO WHILE (T>325.0d0)! Boucle sur les températures
335
336         ALLOCATE(Kmat(Itmax,Cymax))
337         ALLOCATE(Xmat(Itmax,Cymax,dim))
338
339         IF(choix==7) THEN !Diagramme
340
341             IF(side==1) THEN !Calculs côté GAUCHE
342                 IF (z1<0.90d0) THEN
343                     T=T-0.1d0
344                 ELSE
345                     T=T-0.02d0
346                 ENDIF
347
348                 WRITE(*,*) T , " cote 1"
349
350             ELSEIF(side==2) THEN! Calculs côté DROIT
351                 IF (z1>0.001d0 .AND. z1<0.1d0) THEN
352                     T=T-0.05d0
353                 ELSE
354                     T=T-0.02d0
355                 ENDIF
356
357                 WRITE(*,*) T , " cote 2"
358
359             ELSE
360
361             ENDIF
362
363             ! Définition des bornes selon la valeur de z1 (qui change à chaque nouvelle température)
364             LB(1) = epsth !Borne basse de x11
365             UB(1) = z1-epsth !Borne haute de x11
366             LB(2) = z1+epsth !Borne basse de x12
367             UB(2) = 1.0d0-epsth !Borne haute de x12
368
369         ELSE !Autres cas hors diagramme
370
371         ENDIF
372
373     !===== DEBUT DES CYCLES =====
374         DO Cy=1,Cymax
375     !===== ALLOCATION DES TABLEAUX =====
376         !== Général
377         ALLOCATE(X(n,dim)) ! Coordonées des krills
378         ALLOCATE(dX(n,dim)) ! Somme des facteurs densité, food et diffusion pour chaque dimension

```

de chaque krill

```
379     ALLOCATE(Xnew(n,dim)) ! Coordonées du krill mises à jour (à la fin d'une itération)
380     ALLOCATE(Xfinal(dim)) !
381     ALLOCATE(Kvect(n))
382     ALLOCATE(Kibest(n))
383     ALLOCATE(Xibest(n,dim))
384     ALLOCATE(Xfood(n,dim))
385     ALLOCATE(Kibestchap(n))
386     !== Densité
387     ALLOCATE(N1(n,dim))
388     !== Food
389     ALLOCATE(F(n,dim))
390     !== Diffusion
391     ALLOCATE(D(n,dim))
392
393 !===== SOUS-PROGRAMMES
394     CALL position_init(LB,UB,dim,X,n) !Randomisation initiale de la position des krills
395     ! Il est important que ce sous-programme soit séparé des autres,
396     ! Sinon à chaque nouvelle itération, Xi serait randomisé.
397     CALL delta_t(dim,UB,LB,dt)
398
399     ! Calcul des f(x) et du minimal f à l'itération N
400     DO i=1,n
401         Kvect(i) = K(X,n,dim,i,z1,T,P,choix,side,etat)
402     ENDDO
403
404     DO i=1,n
405         Kibest(i)=1.d9
406         DO j=1,dim
407             Xibest(i,j) = X(i,j)
408         ENDDO
409     ENDDO
410
411     Kfood = 1.d9
412     Xfood(1,:)=1.d9
413
414     DO i=1,n
415         DO j=1,dim
416             F(i,j)=0.d0
417             N1(i,j)=0.d0
418             D(i,j)=0.d0
419         ENDDO
420     ENDDO
421
422 !===== POINT DE DEPART DES ITERATIONS =====
423     DO It=1,Itmax
424         CALL densite(n,dim,X,Kvect,It,Itmax,N1) !Sous-programme calculant la partie N de dX/dt
425         CALL food(n,dim,X,Kvect,It,Itmax,F,choix,Kibest,Kfood,Xfood,Kibestchap,z1,T,P,side)
!Sous-programme calculant la partie F de dX/dt
426         CALL diffusion(n,dim,It,Itmax,D) !Sous-programme calculant la partie D de dX/dt
427
428         !== Calcul de dXi et calcul du nouveau X après itération
429         DO i=1,n
430             DO j=1,dim
431                 dX(i,j) = F(i,j) + N1(i,j) + D(i,j)
432                 Xnew(i,j) = X(i,j) + dt*dX(i,j)
433             ENDDO
434         ENDDO
435
436 !===== OPERATEURS GENETIQUES =====
437         CALL crossover(n,dim,Xnew,Kibestchap)
438         CALL mutation(n,dim,Xnew,Kibestchap)
439
440         DO i=1,n
441             DO j=1,dim
442                 ! Il est possible qu'un krill sorte du domaine de définition de la fonction,
```

```

443      ! c'est à dire pour une valeur de X en dehors des bornes définies plus haut.
444      ! La condition est :
445      ! - si la position d'un krill est toujours dans le domaine = la valeur de X est
inchangée
446      ! - si le krill est en dehors du domaine = X est randomisé à nouveau
447      IF(Xnew(i,j)<=LB(j) .OR. Xnew(i,j)>=UB(j)) THEN
448          CALL RANDOM_NUMBER(RanNb)
449          Xnew(i,j) = LB(j) + (UB(j)-LB(j))*RanNb
450
451      ELSE
452
453      ENDIF
454
455      ENDDO
456  ENDDO
457
458  !===== MISE A JOUR DES POSITIONS =====
459      DO i=1,n
460          DO j=1,dim
461              X(i,j) = Xnew(i,j) !On donne à X la valeur de Xnew pour pouvoir effectuer les
itérations suivantes
462          ENDDO
463      ENDDO
464
465      ! Calcul des f(x) et du f minimal à l'itération N : stockage de la valeur dans une
matrice ainsi que ses coordonnées (X)
466      Kmat(It,Cy)=1.d9 !Initialisation par une valeur très élevée dans la matrice, pour être
vite remplacée par une autre
467      DO i=1,n
468          Kvect(i) = K(X,n,dim,i,z1,T,P,choix,side,etat)
469
470          IF(Kmat(It,Cy)>Kvect(i)) THEN
471              Kmat(It,Cy)=Kvect(i)
472              DO j=1,dim
473                  Xmat(It,Cy,j)=X(i,j)
474              ENDDO
475
476          ELSE
477
478          ENDIF
479
480      ENDDO
481
482      valmin=MINVAL(Kvect)
483      valmax=MAXVAL(Kvect)
484
485      sommean=0.d0
486      DO i=1,n
487          sommean=sommean+Kvect(i)
488      ENDDO
489      valmean=sommean/DFLOAT(n)
490
491      IF(It==1) THEN
492          initmin=valmin
493          initmax=valmax
494          initmean=valmean
495      ELSE
496          IF(valmin<initmin) THEN
497              initmin=valmin
498
499          ELSE
500              ENDDO
501              initmax=valmax
502              initmean=valmean
503
504      ENDIF

```

```

505             WRITE(30,*) Cy,It,initmin,initmax,initmean
506
507
508         ENDDO
509 !===== FIN DES ITERATIONS =====
510 !===== DEALLOCATION DES TABLEAUX A CHAQUE CYCLE (APRES LES ITERATIONS) =====
511         !== Général
512         DEALLOCATE(X)
513         DEALLOCATE(dX)
514         DEALLOCATE(Xnew)
515         DEALLOCATE(Xfinal)
516         DEALLOCATE(Kvect)
517         DEALLOCATE(Kibest)
518         DEALLOCATE(Xibest)
519         DEALLOCATE(Xfood)
520         DEALLOCATE(Kibestchap)
521         !== Densité
522         DEALLOCATE(N1)
523         !== Food
524         DEALLOCATE(F)
525         !== Diffusion
526         DEALLOCATE(D)
527         WRITE(*,*) "Cycle no : ",Cy
528
529     ENDDO
530 !===== FIN DES CYCLES =====
531
532     ! Récupération de la meilleure valeur (la valeur la plus basse) parmi toutes les itérations
et les cycles effectués
533     ! On en déduit les coordonnées correspondantes
534     Kfinal=1.d9
535     DO Cy=1,Cymax
536         DO It=1,Itmax
537             IF(Kfinal>Kmat(It,Cy)) THEN
538                 Kfinal=Kmat(It,Cy)
539                 DO j=1,dim
540                     Xfin(1,j)=Xmat(It,Cy,j)
541                 ENDDO
542             ELSE
543                 ENDDIF
544         ENDDO
545     ENDDO
546
547 !===== AFFICHAGE DES RESULTATS =====
548     WRITE(*,*) "=====
549     WRITE(*,*)
550     WRITE(*,*) "Valeur de f minimisee : ",Kfinal ! Valeur de f minimisée
551     WRITE(*,*)
552     DO j=1,dim
553         WRITE(*,*) "Valeur de X ",j," :",Xfin(1,j) ! Coordonées correspondantes
554     ENDDO
555     WRITE(*,*)
556
557     !== Partie thermo ==
558     IF(choix==6) THEN ! Choix 6 : Affichage de l'état physique du point considéré
559         IF(DABS(DABS(Xfin(1,1))-DABS(Xfin(1,2)))<1.d-1) THEN
560             IF(etat==2) THEN
561                 WRITE(*,*) "=====> Monophasique vapeur"
562                 WRITE(*,*)
563             ELSE
564                 WRITE(*,*) "=====> Monophasique liquide"
565                 WRITE(*,*)
566             ENDDIF
567         ELSE
568             WRITE(*,*) "=====> Diphasique"
569             WRITE(*,*)

```

```

570         ENDIF
571
572     ELSEIF(choix==7) THEN ! Choix 7 : Stockage des valeurs dans un fichier
573
574         IF(DABS(Xfin(1,1)-Xfin(1,2))>1.d-3) THEN ! On veut x11 et x12 non égaux
575             IF(side==1) THEN ! Coté gauche
576                 WRITE(20,*) Xfin(1,1),Xfin(1,2),T
577
578                 ELSEIF(side==2) THEN ! Coté droit
579                     ! Avec '1-x', on 'retourne' le diagramme obtenu pour avoir la partie de droite
dans le bon sens
580                     WRITE(20,*) 1.d0-Xfin(1,1),1.d0-Xfin(1,2),T
581
582                 ELSE
583
584                 ENDIF
585             ELSE
586
587             ENDIF
588             z1=DABS(Xfin(1,2)+Xfin(1,1))/2.d0
589         ELSE ! Si choix différent du tracé du diagramme, alors on sort de la boucle des températures
590             EXIT
591         ENDIF
592
593         DEALLOCATE(Kmat)
594         DEALLOCATE(Xmat)
595
596     ENDDO ! Fermeture de la boucle sur les températures
597 ENDDO ! Fermeture de la boucle sur les 'côtés'
598
599 CLOSE(20) ! Fermeture du fichier d'acquisition des données pour le diagramme de phases
600 CLOSE(30)
601 WRITE(*,*)
602 CALL ETIME(Tarray, time)
603 WRITE(*,*)"Temps de calcul : ",time, "s"
604 WRITE(*,*)
605 WRITE(*,*) "Rock'n'roll ! \m/"
606 WRITE(*,*) "'Entree' pour quitter"
607 READ(*,*)
608
609
610 END PROGRAM
611
612 !===== FONCTION OBJECTIF =====
613 ! Ce sous programme regroupe toutes les fonctions minimisables par le KH.
614 ! 7 fonctions objectifs selon le choix de l'utilisateur :
615 !   - Choix 1 : Fonction de Goldstein-Price
616 !   - Choix 2 : Fonction de Camel Back - 6 Hump
617 !   - Choix 3 : Fonction de Ackley
618 !   - Choix 4 : Fonction de Rastrigin
619 !   - Choix 5 : Fonction Quartic
620 !   - Les choix 6 et 7 utilisent la même fonction objectif : énergie de Gibbs d'un mélange
eau+méthyl-tert-butyl-éther
621
622 REAL(8) FUNCTION K(x,n,dim,i,z1,T,P,choix,side,etat)
623     IMPLICIT NONE
624     !Variables communes aux fonctions
625     INTEGER::i,n,dim,j,choix,side,etat
626     REAL(8)::x(n,dim),somme,sommel
627     REAL(8),PARAMETER::pi=3.14159265358979323846d0
628
629     !Variables thermodynamiques
630     REAL(8),PARAMETER::R=8.3144621d0
631     REAL(8)::z1,T,P
632     REAL(8)::tc
633     REAL(8)::b12,b21,alpha,tau12,tau21,G12,G21,gamma1(dim),gamma2(dim)

```

```

634 REAL(8)::A(dim),B(dim),C(dim),Psat(dim)
635 REAL(8)::act11liq,act12liq,act21liq,act22liq,act11gaz,act12gaz,act21gaz,act22gaz,x2(dim)
636 REAL(8)::G1liq,G1gaz,G1,G2liq,G2gaz,G2
637
638 IF(choix==1) THEN ! Choix 1 : Fonction de Goldstein-Price
639
640 K=(1.d0+((x(i,1)+x(i,2)+1.d0)**2)*(19.d0-14.d0*x(i,1)+3.d0*x(i,1)**2-14.d0*x(i,2)+&
641 &6.d0*x(i,1)*x(i,2)+3.d0*x(i,2)**2))*(30.d0+((2.d0*x(i,1)-3.d0*x(i,2))**2)*(18.d0-&
642 &32.d0*x(i,1)+12.d0*x(i,1)**2+48.d0*x(i,2)-36.d0*x(i,1)*x(i,2)+27.d0*x(i,2)**2))
643
644 ELSEIF(choix==2) THEN ! Choix 2 : Fonction de Camel Back - 6 Hump
645
646 K=4.d0*(x(i,1))**2-2.1d0*(x(i,1))**4+((x(i,1))**6)/3.d0+x(i,1)*x(i,2)-&
647 &4.d0*(x(i,2))**2+4.d0*(x(i,2))**4
648
649 ELSEIF(choix==3) THEN ! Choix 3 : Fonction de Ackley
650
651 somme=0.d0
652 sommel=0.d0
653 DO j=1,dim
654     somme=somme+(x(i,j))**2
655     sommel=sommel+DCOS(2.d0*pi*x(i,j))
656 ENDDO
657
658 K=(-20.d0)*DEXP((-0.2d0)*DSQRT(somme/DFLOAT(dim)))-&
659 &DEXP(sommel/DFLOAT(dim))+20.d0+DEXP(1.d0)
660
661 ELSEIF(choix==4) THEN ! Choix 4 : Fonction de Rastrigin
662
663 somme=0.d0
664
665 DO j=1,dim
666     somme=somme+(x(i,j)**2-10.d0*DCOS(2.d0*pi*x(i,j)))
667 ENDDO
668
669 K = 10.d0*DFLOAT(dim) + somme
670
671 ELSEIF(choix==5) THEN ! Choix 5 : Fonction Quartic
672
673 somme=0.d0
674
675 DO j=1,dim
676     somme=somme+DFLOAT(j)*X(i,j)**4
677 ENDDO
678
679 K = somme
680
681 ELSEIF(choix==6 .OR. choix==7) THEN ! Choix 6 et 7 : partie thermo
682
683 tc = T-273.15d0 ! Conversion des degrés Kelvin en degrés Celsius
684
685 IF(side==1) THEN ! Constantes pour la partie de GAUCHE : eau (1) + méthyl-tert-butyl-éther (2)
686
687     ! Valeurs des coefficients de l'equation d'Antoine pour les deux constituants
688     A(2)=5.1792d0
689     A(1)=3.9698d0
690     B(2)=1722.09d0
691     B(1)=1097.67d0
692     C(2)=232.8847d0
693     C(1)=221.8468d0
694
695     ! Parametres du modele NRTL pour le melange méthyl-tert-butyl-éther / eau
696     b21=9203.89d0
697     b12=5704.88d0
698     alpha=0.3d0
699

```

```

700     ELSE ! Constantes pour la partie de DROITE : eau (2) + méthyl-tert-butyl-éther (1)
701         ! Les paramètres sont inversés
702
703         ! Valeurs des coefficients de l'equation d'Antoine pour les deux constituants
704         A(1)=5.1792d0
705         A(2)=3.9698d0
706         B(1)=1722.09d0
707         B(2)=1097.67d0
708         C(1)=232.8847d0
709         C(2)=221.8468d0
710
711         ! Parametres du modele NRTL pour le melange méthyl-tert-butyl-éther / eau
712         b12=9203.89d0
713         b21=5704.88d0
714         alpha=0.3d0
715
716     ENDF
717
718     tau12=(b12/(R*T))
719     tau21=(b21/(R*T))
720     G12=DEXP((-alpha)*tau12)
721     G21=DEXP((-alpha)*tau21)
722
723     ! Calcul des pressions de vapeurs des constituants eau (1) et méthyl-tert-butyl-éther (2) à partir
de l'equation d'Antoine
724     DO j=1,dim
725         Psat(j)=10.d0**(A(j)-(B(j)/(tc+C(j))))
726     ENDDO
727
728     !Calcul des coefficients d'activité phase I et II
729     DO j=1,dim
730         x2(j)=1.d0-x(i,j)
731         gamma1(j)=DEXP((x2(j)**2)*((tau21*(G21/(x(i,j)+(x2(j)*G21)))**2)+((tau12*G12)/(x2(j)+(x(i,j)*
G12))**2)))
732         gamma2(j)=DEXP((x(i,j)**2)*((tau12*(G12/(x2(j)+(x(i,j)*G12)))**2)+((tau21*G21)/(x(i,j)+(x2(j)*
G21))**2)))
733     ENDDO
734
735     !Calcul des activités du constituant 1
736     IF (P>Psat(1)) THEN
737         act11liq=x(i,1)*gamma1(1)
738         act11gaz=(P*x(i,1))/Psat(1)
739         act12liq=x(i,2)*gamma1(2)
740         act12gaz=(P*x(i,2))/Psat(1)
741     ELSE
742         act11liq=(Psat(1)*x(i,1)*gamma1(1))/P
743         act11gaz=x(i,1)
744         act12liq=(Psat(1)*x(i,2)*gamma1(2))/P
745         act12gaz=x(i,2)
746     ENDF
747
748     ! Calcul des activités du constituant 2
749     IF (P>Psat(2)) THEN
750         act21liq=(1.d0-x(i,1))*gamma2(1)
751         act21gaz=(P*(1.d0-x(i,1)))/Psat(2)
752         act22liq=(1.d0-x(i,2))*gamma2(2)
753         act22gaz=(P*(1.d0-x(i,2)))/Psat(2)
754     ELSE
755         act21liq=(Psat(2)*(1.d0-x(i,1))*gamma2(1))/P
756         act21gaz=(1.d0-x(i,1))
757         act22liq=(Psat(2)*(1.d0-x(i,2))*gamma2(2))/P
758         act22gaz=(1.d0-x(i,2))
759     ENDF
760
761     !Calcul de Gibbs molaire pour la phase 1 (liquide ou gaz)
762     G1liq=(x(i,1)*DLOG(act11liq)+(1.d0-x(i,1))*DLOG(act21liq))

```

```

763      G1gaz=(x(i,1)*DLOG(act11gaz)+(1.d0-x(i,1))*DLOG(act21gaz))
764
765      G1=MIN(G1liq,G1gaz)
766
767      IF (DABS(DABS(G1)-DABS(G1liq))<=1.d-10) THEN
768          etat=1
769      ELSE
770          etat=2
771      ENDIF
772
773      ! Calcul de Gibbs molaire pour la phase 2 (liquide ou gaz)
774      G2liq=(x(i,2)*DLOG(act12liq)+(1.d0-x(i,2))*DLOG(act22liq))
775      G2gaz=(x(i,2)*DLOG(act12gaz)+(1.d0-x(i,2))*DLOG(act22gaz))
776
777      G2=MIN(G2liq,G2gaz)
778
779      ! Calcul de Gibbs molaire totale du mélange
780      K=((x(i,2)-z1)*G1)/(x(i,2)-x(i,1))+((z1-x(i,1))*G2)/(x(i,2)-x(i,1))
781
782      ELSE
783
784      ENDIF
785
786  END FUNCTION
787
788  !===== RANDOMISATION DE LA POSITION DES KRILLS = ETAT INITIAL
789  ! Sous programme appelé pour randomiser la position d'un krill
790  ! Il est appelé une seule fois au début de l'algorithme KH pour initialiser la position des krills
791  ! La randomisation permet de générer une population aléatoire dans les bornes définies de la fonctions
objectif
792
793  SUBROUTINE position_init(b,a,dim,X,n)
794      IMPLICIT NONE
795      INTEGER::i,j,n,dim
796      REAL(8)::rand(n,dim),a(dim),b(dim),X(n,dim)
797
798      DO i=1,n
799          DO j=1,dim
800              CALL RANDOM_NUMBER(rand(i,j))
801          ENDDO
802      ENDDO
803
804      DO i=1,n
805          DO j=1,dim
806              ! 'a(j)' et 'b(j)' sont les bornes basse et haute de la variable j
807              X(i,j) = a(j) + (b(j)-a(j))*rand(i,j)
808          ENDDO
809      ENDDO
810
811  END SUBROUTINE
812
813  !=====
814
815          !===== PARTIE DENSITE (N) =====
816
817  !=====
818
819  ! Effet global d'attraction des krills entre eux afin augmenter la densité en krills de l'espace vers une
valeur minimale
820
821  !===== DETERMINATION DES VOISINS POUR CHAQUE KRILL
822  ! Le but principal de ce sous programme est de pouvoir déterminer à chaque itération :
823  ! - la distance qui sépare un krill des autres krills
824  ! - et de déduire si un krill vis à vis d'un autre krill est un voisin ou non
825  SUBROUTINE neighbor(n,dim,X,Xdiff,nbor,Xnorm)
826      IMPLICIT NONE

```

```

827     INTEGER::i,j,h,n,dim
828     INTEGER::NN(n)
829     REAL(8)::som_Xnorm(n),ds(n)
830     REAL(8)::Xdifff(n,n,dim),Xnorm(n,n),X(n,dim),som_carre(n,n)
831     LOGICAL::nbor(n,n)
832
833     DO i=1,n
834         DO h=1,n
835             IF(h>=i) THEN
836                 !'h=i' concerne le même krill, le calcul de la différence vaut automatiquement zéro
837                 DO j=1,dim
838                     Xdifff(i,h,j) = 0.d0
839                 ENDDO
840                 !On définit 'Xnorm = ||Xh-Xi||' la norme de la différence des vecteurs position des krills
'h' et 'i'
841                 !Cette différence de vecteurs donne un nouveau vecteur appelé ici 'Xdifff'
842                 Xnorm(i,h)=0.d0 !Le calcul de la norme vaut également zéro dans cette situation
843
844             ELSE
845                 DO j=1,dim !Boucle sur les dimensions
846                 !Calcul de 'Xh-Xi' : soustraction de matrices de dimension (n x dim) => donne une
matrice de même dimension
847                 !Un krill peut posséder 1 ou plusieurs voisins, c'est à dire que pour un krill 'i'
donné, il existe plusieurs 'Xh-Xi'
848                 !Il faut alors ajouter une dimension à Xdifff pour prendre en compte ces voisins, d'où
'Xdifff(i,h,j)'
849                 Xdifff(i,h,j) = X(h,j) - X(i,j)
850             ENDDO
851             !Calcul de ||Xh-Xi|| : pour des krills 'i' et 'h' donnés, on aura une valeur.
852             !De la même manière que précédemment, il faut prendre en compte les voisins de chaque krill
853             !On aura alors une matrice de dimension (n x n)
854             !La différence avec 'Xh-Xi' vient du calcul de la norme qui 'enlève' une dimension à la
matrice
855
856             som_carre(i,h) = 0.d0! Ici, on calcule la norme des positions
857             DO j=1,dim
858                 som_carre(i,h) = som_carre(i,h) + (Xdifff(i,h,j))**2
859             ENDDO
860
861             Xnorm(i,h) = DSQRT(som_carre(i,h))
862
863         ENDDIF
864     ENDDO
865 ENDDO
866
867 som_Xnorm(:)=0.d0
868 DO i=1,n
869     DO h=1,n
870         som_Xnorm(i) = som_Xnorm(i) + Xnorm(i,h) !Calcul de la somme des normes pour le calcul de
'ds,i'
871     ENDDO
872 ENDDO
873
874 !Calcul de la 'sensing distance' pour chaque krill, elle permettra de déterminer les voisins de chacun
875 !C'est un vecteur, il y a une seule distance pour chaque krill
876 DO i=1,n
877     ds(i) = (1.d0/(5.d0*DFLOAT(n)))*som_Xnorm(i)
878 ENDDO
879
880 !Grâce à la 'sensing distance' précédemment calculée :
881 ! attribution d'une valeur booléenne à chaque krill 'h' pour un krill 'i'
882 ! 'TRUE' : c'est un voisin
883 ! 'FALSE' : ce n'est pas un voisin
884 DO i=1,n
885     NN(i)=0
886     DO h=1,n

```

```

887     IF (i==h) THEN
888         !Le krill 'h' est le meme que le krill 'i'
889         !Comme un krill ne peut être voisin de lui-même => FALSE
890         nbor(i,h) = .FALSE.
891     ELSE
892         IF (Xnorm(i,h)<ds(i)) THEN
893             !Le krill h EST un voisin => TRUE
894             nbor(i,h) = .TRUE.
895             NN(i) = NN(i) + 1
896         ELSE
897             !Le krill h N'EST PAS un voisin => FALSE
898             nbor(i,h) = .FALSE.
899         ENDIF
900     ENDIF
901 ENDDO
902 ENDDO
903
904 END SUBROUTINE
905
906
907 !===== CALCUL DE ALPHA LOCAL
908 ! Les calculs d'Alpha local ne s'appliquent qu'aux krills voisins entre eux.
909 ! C'est une composante d'attraction des individus entre eux.
910 SUBROUTINE alpha_local(n,dim,X,Kvect,Alocal,Km,Xm,nbor)
911     IMPLICIT NONE
912     INTEGER::i,j,h,n,dim
913     REAL(8)::Kmax,Kmin
914     REAL(8)::Alocal(n,dim),Km(n,n),Xm(n,n,dim),Xdifff(n,n,dim),X(n,dim),Xnorm(n,n),Kvect(n)
915     REAL(8),PARAMETER::epsi=1.d-10 !'Epsilon' : évite les singularités, il doit être petit et positif
916     LOGICAL::nbor(n,n)
917
918     CALL neighbor(n,dim,X,Xdifff,nbor,Xnorm)
919
920     ! Calcul de toutes les valeurs de Ki en fonction de Xi
921     ! Cette étape permet par la suite de sélectionner la pire et la meilleure valeur de K
922     ! Kmax étant la valeur la plus haute donc la moins bonne
923     ! Kmin la valeur la plus basse, donc la meilleure
924
925     Kmax = MAXVAL(Kvect)
926     Kmin = MINVAL(Kvect)
927
928     DO i=1,n
929         DO h=1,n
930             IF(nbor(i,h) .EQV. .TRUE.) THEN
931                 !On ne calculera que les Xm et Km des voisins
932                 DO j=1,dim
933                     !Xdifff a pour dimension (n x n x dim), Xnorm (n x n),
934                     !On obtient alors un Xm de dimension (n x n x dim)
935                     Xm(i,h,j) = (Xdifff(i,h,j))/(Xnorm(i,h) + epsi)
936                 ENDDO
937                 !Calcul de 'Ki-Kh' : K est la valeur renvoyée par la fonction objectif f(x1..xdim)
938                 !De la même manière que '||Xh-Xi||', on aurait un vecteur si chaque krill n'avait qu'un
seul voisin
939                 !Or, il peut y en avoir plusieurs, donc on obtient une matrice (n x n)
940                 !Comme Kmax et Kmin sont des valeurs fixes, on a alors Km de dimension (n x n)
941                 Km(i,h) = (Kvect(i) - Kvect(h))/(Kmax-Kmin)
942             ELSE
943                 !Attribuer la valeur 0 au reste, permet au produit Km*Xm de rester nul
944                 !et de ne pas avoir de poids dans la somme des produits pour alpha local ('Alocal')
945                 DO j=1,dim
946                     Xm(i,h,j) = 0.d0
947                     Km(i,h) = 0.d0
948                 ENDDO
949             ENDIF
950         ENDDO
951     ENDDO

```

```

952
953     DO i=1,n
954         DO j=1,dim
955             DO h=1,n
956                 !Le produit d'un Km avec un Xm donne une matrice (n x n x dim)
957                 !La somme des Km*Xm est la somme, pour un krill 'i' donné, de tous les produits possibles
avec ses voisins.
958                 !Si on fesait la somme des Km on aurait un vecteur de dimension (n)
959                 !Xm comporte une dimension supplémentaire par rapport à Km : les variables.
960                 !La somme des Xm renverrait donc une matrice (n x dim)
961                 !La dimension de Alpha local est donc bien une matrice (n x dim)
962                 !Encore une fois, cette somme par rapport aux voisins permet d'"enlever" une dimension des
matrices du calcul.
963                 Alocal(i,j) = Alocal(i,j) + Km(i,h)*Xm(i,h,j)
964             ENDDO
965         ENDDO
966     ENDDO
967
968 END SUBROUTINE
969
970 !===== CALCUL DE ALPHA TARGET
971 ! On repère le krill ayant la meilleure position dans la population (valeur la plus basse de la fonction
objectif)
972 ! pour en faire une cible pour la population : ce facteur permet un mouvement global de la population vers
un objectif
973 SUBROUTINE alpha_target(n,dim,Itmax,It,X,Kvect,Atarget,Alocal)
974     IMPLICIT NONE
975     INTEGER::It,Itmax,n,dim,i,j,h
976     INTEGER::locX1(n,dim),locX2(n)
977     REAL(8)::Km(n,n),Xm(n,n,dim),Alocal(n,dim),X(n,dim),Kibest(n),Atarget(n,dim),Kvect(n)
978     REAL(8)::Cbest,RanNb
979     LOGICAL::nbor(n,n)
980
981     !Génération d'un nombre aléatoire pour le calcul de Cbest
982     !Ce nombre permet l'exploration
983     CALL RANDOM_NUMBER(RanNb)
984
985     !Cbest est dépendant des itérations : une seule valeur
986     Cbest = 2.d0*(RanNb + DFLOAT(It)/DFLOAT(Itmax))
987
988     !L'appel du sous-programme 'alpha-local' permet entre autres de récupérer les valeurs de Km et Xm
989     CALL alpha_local(n,dim,X,Kvect,Alocal,Km,Xm,nbor)
990
991     DO i=1,n
992
993         !'Kibest' est un vecteur (n)
994         Kibest(i) = MINVAL(DABS(Km(i,:)))
995
996         !Il faut localiser le krill 'h' pour chaque krill 'i' dans la matrice pour déduire ses coordonnées
997         !'Xibest' prend en compte les différentes variables de chaque krill c'est donc une matrice (n x
dim)
998         locX1(i,1:dim) = MINLOC(DABS(Km(i,:)))
999
1000         !La fonction MINLOC n'est pas très pratique car elle renvoie un vecteur ligne (même si une seule
valeur est renvoyée ici)
1001         !Il faut donc convertir la matrice locX1 obtenue en un vecteur colonne locX2
1002         locX2(i)=locX1(i,1)
1003
1004         DO j=1,dim
1005             ! L'écriture 'X(i,locX2(i),j)' n'est pas valable en Fortran 90
1006             ! Il faut passer par une valeur ponctuelle 'h' qui varie à chaque boucle
1007             h=locX2(i)
1008             Atarget(i,j) = Cbest*Kibest(i)*Xm(i,h,j)
1009         ENDDO
1010     ENDDO
1011

```

```

1012 END SUBROUTINE
1013
1014 !===== CALCUL DE ALPHA
1015 ! Caclul d'alpha par simple somme des effets d'alpha target et local
1016 SUBROUTINE alpha_global(n,dim,X,Kvect,It,Itmax,Aglobal)
1017   IMPLICIT NONE
1018   INTEGER::n,dim,i,j,Itmax,It
1019   REAL(8)::Alocal(n,dim),Atarget(n,dim),Aglobal(n,dim),X(n,dim),Kvect(n)
1020
1021   CALL alpha_target(n,dim,Itmax,It,X,Kvect,Atarget,Alocal)
1022
1023   ! Le calcul d'alpha est uniquement la somme d'alpha target avec alpha local
1024   DO i=1,n
1025     DO j=1,dim
1026       Aglobal(i,j) = Atarget(i,j) + Alocal(i,j)
1027     ENDDO
1028   ENDDO
1029
1030 END SUBROUTINE
1031
1032 !===== CALCUL DE N
1033 ! Le calcul du facteur densité dépend d'alpha et de sa précédente valeur.
1034 ! A l'itération 1, cette valeur est initialisée dans le programme principal
1035 SUBROUTINE densite(n,dim,X,Kvect,It,Itmax,N0)
1036   IMPLICIT NONE
1037   INTEGER::It,Itmax,n,dim,i,j
1038   REAL(8)::N1(n,dim),N0(n,dim),Aglobal(n,dim),X(n,dim),Kvect(n)
1039   REAL(8)::Wn
1040   REAL(8),PARAMETER::Nmax = 0.01d0 !Vitesse max des krills donnée par la publication
1041
1042   !N1 et N0, comme X, sont des matrices de dimension (n,dim)
1043   !Wn est une valeur dépendante uniquement des itérations ==> pas un vecteur ni une matrice
1044
1045   CALL alpha_global(n,dim,X,Kvect,It,Itmax,Aglobal)
1046   Wn = 0.1d0 + 0.8d0*(1.d0-DFLOAT(It)/DFLOAT(Itmax))
1047
1048   DO i=1,n
1049     DO j=1,dim
1050       N1(i,j) = Nmax*Aglobal(i,j) + Wn*N0(i,j)
1051       N0(i,j) = N1(i,j)
1052     ENDDO
1053   ENDDO
1054
1055 END SUBROUTINE
1056
1057 !=====
1058
1059           !===== PARTIE FOOD (F) =====
1060
1061 !=====
1062
1063 ! L'objectif principal de ce facteur est de créer un point de nourriture (ici une valeur minimale globale)
1064 ! qui servira d'objectif dans le déplacement des krills.
1065
1066 !===== CALCUL DES BETA
1067 ! Dans beta total, 2 effets combinés :
1068 ! - L'attraction des krills vers la nourriture
1069 ! - L'attraction des krills vers leur meilleure valeur obtenue précédement
1070 SUBROUTINE beta_total(n,dim,X,Kvect,It,Itmax,betafood,choix,Kibest,Kfood,Xfood,Kibestchap,zl,T,P,side)
1071   IMPLICIT NONE
1072   REAL(8)::K,Kmax,Kmin,Cfood,Kfood,Kfoodnew,zl,T,P !Scalaires
1073   REAL(8)::somX(dim),somK(dim),Xfood(n,dim),Xnormfood(n),som_carre(n),Kdiff(n)
1074   REAL(8)::Kifood(n),Kibest(n),Kvect(n),Xnorm(n),Kibestchap(n),Xfoodnew(n,dim) ! Vecteurs
1075   REAL(8)::X(n,dim),Xifood(n,dim),Xdif(n,dim),betafood(n,dim),betabest(n,dim)
1076   REAL(8)::beta(n,dim),Xibest(n,dim),Xichap(n,dim) ! Tableaux/Matrices 2x2
1077   REAL(8),PARAMETER::eps=1.d-10 !Constantes

```

```

1078     INTEGER::n,dim,i,j,h,Itmax,It,choix,side,etat !Entiers
1079
1080 !===== CALCUL DE BETA FOOD =====
1081
1082 !Initialisation des sommes, pour le calcul des parties de Xfood
1083 somX(:)=0.d0
1084 somK(:)=0.d0
1085
1086 DO j=1,dim
1087     DO i=1,n
1088         somX(j)=somX(j)+X(i,j)/Kvect(i)
1089         somK(j)=somK(j)+1.d0/Kvect(i)
1090     ENDDO
1091 ENDDO
1092
1093 DO j=1,dim
1094     Xfoodnew(1,j)=somX(j)/somK(j)
1095     ! 'Xfoodnew' est la nouvelle valeur calculée de Xfood que l'on comparera à l'ancienne
1096 ENDDO
1097
1098
1099 !Calcul de F(Xfood)
1100 Kfoodnew = K(Xfoodnew,n,dim,1,z1,T,P,choix,side,etat)
1101
1102 ! Si un krill possède une meilleure position que Xfood (un f(x) plus bas)
1103 ! alors on définit la position de ce krill en tant que nouveau Xfood
1104 DO i=1,n ! On teste toutes les valeurs de f à Xfood
1105     IF(Kfoodnew>Kvect(i)) THEN
1106         Kfoodnew = Kvect(i)
1107         DO h=1,dim
1108             ! On récupère les coordonnées du krill optimal
1109             Xfoodnew(1,h) = X(i,h)
1110         ENDDO
1111     ELSE
1112     ENDF
1113 ENDDO
1114
1115 ! Si la nouvelle valeur obtenue de Xfood est meilleure que l'ancienne, on la garde et on récupère ses
1116 coordonnées
1117 ! Dans le cas contraire, on garde l'ancien Xfood
1118 IF(Kfoodnew<Kfood) THEN
1119     Kfood = Kfoodnew
1120     DO j=1,dim
1121         Xfood(1,j)=Xfoodnew(1,j)
1122     ENDDO
1123 ELSE
1124 ENDF
1125
1126 !=== Calcul de X^ifood ===
1127 DO i=1,n
1128     DO j=1,dim
1129         Xdiff(i,j)=Xfood(1,j)-X(i,j)
1130     ENDDO
1131 ENDDO
1132
1133 som_carre(:) = 0.d0
1134 DO i=1,n
1135     DO j=1,dim
1136         som_carre(i) = som_carre(i) + (Xdiff(i,j))**2
1137     ENDDO
1138 ENDDO
1139
1140 DO i=1,n
1141     !Calcul de la norme de la différence des deux vecteurs
1142     Xnormfood(i) = DSQRT(som_carre(i))
1143     DO j=1,dim

```

```

1143         Xifood(i,j)=Xdifff(i,j)/(Xnormfood(i)+eps)
1144     ENDDO
1145 ENDDO
1146
1147 !=== Calcul de K^ifood ===
1148 DO i=1,n
1149     Kdiffer(i) = Kvect(i) - Kfood
1150 ENDDO
1151
1152 Kmax = MAXVAL(Kvect) ! Kworst : valeur maximale de f calculée
1153 Kmin = MINVAL(Kvect) ! Kbest : valeur minimale de f calculée
1154
1155 DO i=1,n
1156     Kifood(i) = Kdiffer(i)/(Kmax-Kmin)
1157 ENDDO
1158
1159 ! Le calcul de Cfood est guidé dans la publication
1160 Cfood=2.d0*(1.d0-(DFLOAT(It)/DFLOAT(Itmax)))
1161
1162 !=== Calcul de beta food ===
1163 DO i=1,n
1164     DO j=1,dim
1165         betafood(i,j) = Cfood*Kifood(i)*Xifood(i,j) !Calcul de beta food
1166     ENDDO
1167 ENDDO
1168
1169 !===== FIN DE CALCUL DE BETA FOOD =====
1170
1171 !===== CALCUL DE BETA BEST =====
1172
1173 DO i=1,n
1174     ! Si la nouvelle valeur de f pour un krill est la meilleure parmi toutes les précédentes, on la
garde
1175     ! et betabest = 0 --> on ne fait pas bouger le krill.
1176     ! Si ce n'est pas la meilleure alors on calcule betabest avec son actuelle position et sa
meilleure position obtenue avant
1177     IF(Kvect(i)<Kibest(i)) THEN
1178         betabest(i,:) = 0.d0
1179         Kibest(i) = Kvect(i)
1180         Xibest(i,:) = X(i,:)
1181     ELSE
1182         Kdiffer(i) = Kvect(i) - Kibest(i)
1183         Kibestchapp(i) = Kdiffer(i)/(Kmax-Kmin)
1184
1185         DO j=1,dim
1186             Xdiffer(i,j)=Xibest(i,j)-X(i,j)
1187         ENDDO
1188
1189         !Calcul de la norme de la différence des deux vecteurs
1190         som_carre(i) = 0.d0
1191         DO j=1,dim
1192             som_carre(i) = som_carre(i) + (Xdiffer(i,j))**2
1193         ENDDO
1194
1195         Xnorm(i) = DSQRT(som_carre(i))
1196         DO j=1,dim
1197             Xichap(i,j)=Xdiffer(i,j)/(Xnorm(i)+eps)
1198         ENDDO
1199
1200         DO j=1,dim
1201             betabest(i,j) = Kibestchapp(i)*Xichap(i,j)! Calcul de beta best
1202         ENDDO
1203     ENDDO
1204 ENDDO
1205 ENDDO
1206

```

```

1207 !===== FIN DE CALCUL DE BETA BEST =====
1208
1209 !===== CALCUL DE BETA GLOBAL =====
1210
1211     DO i=1,n
1212         DO j=1,dim
1213             beta(i,j) = betabest(i,j) + betafood(i,j)
1214         ENDDO
1215     ENDDO
1216
1217 !===== FIN DE CALCUL DE BETA GLOBAL =====
1218 END SUBROUTINE
1219
1220 !===== CALCUL DU FOOD
1221 ! Tout comme la densité, l'effet de la nourriture sur le comportement des krill dépend de la précédente
valeur de ce facteur
1222 ! Ici, seul le calcul final du facteur est réalisé
1223 SUBROUTINE food(n,dim,X,Kvect,It,Itmax,F0,choix,Kibest,Kfood,Xfood,Kibestchap,z1,T,P,side)
1224     IMPLICIT NONE
1225     REAL(8)::wf,vf,Kfood,z1,T,P
1226     REAL(8)::F0(n,dim),F1(n,dim),beta(n,dim),X(n,dim),Kvect(n),Kibest(n),Xfood(n,dim),Kibestchap(n)
1227     INTEGER::i,j,n,dim,It,Itmax,choix,side
1228
1229     CALL RANDOM_NUMBER (wf) ! Déjà compris dans l'intervalle [0,1]
1230     vf=2.d-2
1231
1232     CALL beta_total(n,dim,X,Kvect,It,Itmax,beta,choix,Kibest,Kfood,Xfood,Kibestchap,z1,T,P,side)
1233
1234     DO i=1,n
1235         DO j=1,dim
1236             F1(i,j) = vf*beta(i,j) + wf*F0(i,j)
1237             F0(i,j) = F1(i,j)
1238         ENDDO
1239     ENDDO
1240
1241 END SUBROUTINE
1242
1243 !=====
1244
1245             !===== PARTIE DIFFUSION (D) =====
1246
1247 !=====
1248
1249 ! La diffusion est un facteur qui permet de caractériser le comportement aléatoire des krills.
1250 ! Il est composé de 2 nombres aléatoires pour chaque variable de chaque krill
1251 SUBROUTINE diffusion(n,dim,It,Itmax,D)
1252     IMPLICIT NONE
1253     REAL(8)::D(n,dim),Dmax(n),Dmaxnbr(n),delta(n,dim),deltanbr(n,dim)
1254     REAL(8),PARAMETER::aDmax= 2.d-3,bDmax=1.d-2,adelta=-1.d0,bdelta=1.d0
1255     INTEGER::n,dim,i,j,Itmax,It
1256
1257     DO i=1,n
1258         CALL RANDOM_NUMBER(Dmaxnbr)
1259     ENDDO
1260
1261     DO i=1,n
1262         ! Dmax est propre à chaque krill et doit être généré qu'une fois
1263         ! Dmax est un vecteur
1264         Dmax(i) = aDmax + (bDmax-aDmax)*Dmaxnbr(i)
1265     ENDDO
1266
1267     DO i=1,n
1268         DO j=1,dim
1269             CALL RANDOM_NUMBER(deltanbr(i,j))
1270         ENDDO
1271     ENDDO

```

```

1272
1273     DO i=1,n
1274         DO j=1,dim
1275             ! delta doit être différent pour chaque dimension de chaque krill
1276             ! c'est donc un tableau/matrice
1277             delta(i,j) = adelta + (bdelta-adelta)*deltanbr(i,j)
1278         ENDDO
1279     ENDDO
1280
1281     DO i=1,n
1282         DO j=1,dim
1283             D(i,j) = Dmax(i)*(1.d0 - DFLOAT(It)/DFLOAT(Itmax))*delta(i,j)
1284         ENDDO
1285     ENDDO
1286
1287 END SUBROUTINE
1288
1289 !=====
1290
1291             !===== PARTIE MOTION PROCESS (delta t) =====
1292
1293 !=====
1294
1295 ! Le motion process est calcul général en fin d'itération qui permet d'obtenir la nouvelle position des
1296 ! krills sous l'effet des 3 facteurs (N,F,D).
1297 ! Cet effet de ces facteurs est 'régulé' par un nombre aléatoire qui détermine alors l'impact de ceux ci
1298 ! sur le mouvement du krill
1299 ! Ce qui a pour conséquence de rendre les krills plus ou moins prudents dans leurs mouvements.
1300 ! Ce comportement change à chaque lancement de l'algorithme
1301
1302 SUBROUTINE delta_t(dim,UB,LB,dt)
1303     IMPLICIT NONE
1304     INTEGER::j,dim
1305     REAL(8)::sum,Ct,dt,RanNb
1306     REAL(8)::UB(dim),LB(dim)
1307     REAL(8),PARAMETER::a=0.d0,b=2.d0
1308
1309     !Calcul de la somme des différences sur les bornes des variables
1310     !Cette différence détermine l'espace de recherche
1311     sum = 0.d0 !Initialisation
1312     DO j=1,dim
1313         sum = sum + (UB(j) - LB(j))
1314     ENDDO
1315
1316     !'Ct' est un nombre aléatoire constant entre 0 et 2
1317     !Une valeur basse de 'Ct' rend le krill prudent dans sa recherche
1318     CALL RANDOM_NUMBER(RanNb)
1319     Ct = a + (b-a)*RanNb
1320
1321     dt = Ct*sum
1322
1323 END SUBROUTINE
1324
1325 !=====
1326
1327             !===== PARTIE CROSSOVER =====
1328
1329 ! Le crossover est un opérateur génétique basé sur la probabilité d'un krill à adopter la même position
1330 ! d'un autre krill au hasard.
1331 ! Cette probabilité augmente plus le krill s'éloigne de l'objectif (valeur minimale globale).
1332
1333 SUBROUTINE crossover(n,dim,X,Kibest)
1334     IMPLICIT NONE
1335     INTEGER::n,dim,i,j,r
1336     REAL(8)::Cr,Kibest(n),X(n,dim),rand

```

```

1335     DO i=1,n
1336         Cr = 0.2d0*Kibest(i)
1337         DO j=1,dim
1338
1339             CALL RANDOM_NUMBER(rand)
1340             r=ceiling(rand*DFLOAT(n))
1341
1342             IF(Cr>rand .AND. r/=i) THEN
1343                 X(i,j)=X(r,j)
1344
1345             ELSEIF (Cr>rand .AND. r==i) THEN
1346                 r=ceiling(rand*DFLOAT(n/2))
1347                 X(i,j)=X(r,j)
1348
1349             ELSE
1350                 ENDF
1351
1352         ENDDO
1353     ENDDO
1354
1355 END SUBROUTINE
1356
1357 !=====
1358
1359             !===== PARTIE MUTATION =====
1360
1361 !=====
1362
1363 ! La mutation est un opérateur génétique basé sur la probabilité que la position de 2 krills au hasard
1364 ! influencent le mouvement d'un 3ème.
1365 ! Cette probabilité augmente plus le krill s'éloigne de l'objectif (valeur minimale globale).
1366 SUBROUTINE mutation(n,dim,X,Kibest)
1367     IMPLICIT NONE
1368     INTEGER::n,dim,i,j,p,q,r
1369     REAL(8)::Mu,Kibest(n),X(n,dim),rand1,rand2,Kibestglobal
1370
1371     Kibestglobal=1.d9
1372     DO i=1,n
1373         IF(DABS(Kibest(i))<DABS(Kibestglobal)) THEN
1374             r=i
1375         ELSE
1376             ENDF
1377         ENDDO
1378     DO i=1,n
1379         Mu = 0.05d0*Kibest(i)
1380         !Dans la publication, les auteurs utilisent 'Mu = 0.05d0/Kibest(i)',
1381         !La probabilité de mutation augmente plus le krill approche du point optimal, contrairement à ce
1382         !qui est mentionné.
1383         !Il est préférable d'utiliser la formule ci-dessus si l'on veut respecter la logique de
1384         !l'algorithme.
1385         !A l'aide de diagrammes comparatifs, il a été montré que la mutation est plus pertinente dans ce
1386         !sens.
1387         DO j=1,dim
1388
1389             CALL RANDOM_NUMBER(rand1)
1390             p=ceiling(rand1*DFLOAT(n))
1391
1392             CALL RANDOM_NUMBER(rand2)
1393             q=ceiling(rand2*DFLOAT(n))
1394
1395             IF(Mu>rand1 .AND. p/=i .AND. q/=i) THEN
1396                 X(i,j)=X(i,j)+rand1*(X(p,j)-X(q,j))
1397
1398             ELSEIF(Mu>rand1 .AND. (p==i .OR. q==i)) THEN
1399                 p=ceiling(rand1*DFLOAT(n/2))

```

```
1397         q=ceiling(rand2*DFLOAT(n/2))
1398         X(i,j)=X(i,j)+rand1*(X(p,j)-X(q,j))
1399
1400         ELSE
1401         ENDIF
1402
1403     ENDDO
1404 ENDDO
1405
1406
1407 END SUBROUTINE
1408
1409 !===== FIN DU KH ===== \m/
```

Bibliographie

- [1] X. Hu and R. Eberhart. Solving constrained nonlinear optimization problems with particle swarm optimization. In *Proceedings of the sixth world multiconference on systemics, cybernetics and informatics*, volume 5, pages 203–206. Citeseer, 2002.
- [2] T. Sousa, A. Silva, and A. Neves. Particle swarm based data mining algorithms for classification tasks. *Parallel computing*, 30(5-6):767–783, 2004.
- [3] N. Arana-Daniel, C. Lopez-Franco, and A.Y. Alanis. *Bio-inspired Algorithms for Engineering*. Elsevier Science, Guadalajara, Jalisco, Mexico, 2018.
- [4] R. Eberhart and J. Kennedy. A new optimizer using particle swarm theory. In *Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pages 39–43. IEEE, 1995.
- [5] M. Dorigo, V. Maniezzo, and A. Coloni. Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(1):29–41, 1996.
- [6] C. Blum. *Theoretical and Practical Aspects of Ant Colony Optimization*. Dissertationen zur künstlichen Intelligenz. Aka, Berlin, Germany, 2004.
- [7] N. Nedjah and L. de Macedo Mourelle. *Systems Engineering Using Swarm Particle Optimisation*. Intelligent system engineering book series. Nova Science Publishers, Rio de Janeiro, Rio de Janeiro, Brazil, 2007.
- [8] P.I. Cowling and P. Merz. *Evolutionary Computation in Combinatorial Optimization: 10th European Conference*. LNCS sublibrary: SL 1-theoretical computer science and general issues. Springer, Istanbul, Turkey, 2010.
- [9] E. Elbeltagi, T. Hegazy, and D. Grierson. Comparison among five evolutionary-based optimization algorithms. *Advanced engineering informatics*, 19(1):43–53, 2005.
- [10] C. Solnon. *Ant Colony Optimization and Constraint Programming*. ISTE. Wiley, Hoboken, NJ, USA, 2010.
- [11] F. Xhafa and A. Abraham. *Metaheuristics for Scheduling in Industrial and Manufacturing Applications*. Studies in Computational Intelligence. Springer Berlin, Heidelberg, Germany, 2008.
- [12] R. Kowalczyk. *Computational Collective Intelligence. Semantic Web, Social Networks and Multi-agent Systems: First International Conference*. Lecture Notes in Computer Science. Springer Berlin Heidelberg, Wroclaw, Poland, 2009.
- [13] P.A. Miranda and R.A. Garrido. Incorporating inventory control decisions into a strategic distribution network design model with stochastic demand. *Transportation Research Part E: Logistics and Transportation Review*, 40(3):183–207, 2004.

- [14] D. Karaboga. An idea based on honey bee swarm for numerical optimization. Technical report, Technical report-tr06, Erciyes university, engineering faculty, computer engineering department, 2005.
- [15] A. Rajasekhar, S. Das, and S. Das. μ abc: A micro artificial bee colony algorithm for large scale global optimization. In *Proceedings of the 14th Annual Conference Companion on Genetic and Evolutionary Computation*, GECCO '12, pages 1399–1400, New York, NY, USA, 2012. ACM.
- [16] A.R. Ravindran, R. Ufuk Bilsel, V. Wadhwa, and T. Yang. Risk adjusted multicriteria supplier selection models with applications. *International Journal of Production Research*, 48(2):405–424, 2010.
- [17] B. Rangaiah. The influence of emotional intelligence on life satisfaction and job performance of employees. *Asian Journal of Development Matters*, 4(2):177–183, 2010.
- [18] C.A. Floudas. *Deterministic global optimization: theory, methods and applications*, volume 37. Springer Science & Business Media, Heidelberg, Germany, 2013.
- [19] K.M. Passino. Biomimicry of bacterial foraging for distributed optimization and control. *IEEE control systems*, 22(3):52–67, 2002.
- [20] N.E. Karkalos, A.P. Markopoulos, and J.P. Davim. *Computational Methods for Application in Industry 4.0*. SpringerBriefs in Applied Sciences and Technology. Springer International Publishing, Cham, Switzerland, 2018.
- [21] X. Yang and X. He. Firefly algorithm: recent advances and applications. *International Journal of Swarm Intelligence*, 1(1):36–50, 2013.
- [22] X. Yang. A new metaheuristic bat-inspired algorithm. In *Nature inspired cooperative strategies for optimization (NICSO 2010)*, pages 65–74. Springer Berlin, Heidelberg, Germany, 2010.
- [23] F. van den Bergh and A.P. Engelbrecht. A new locally convergent particle swarm optimiser. In *Systems, Man and Cybernetics, 2002 IEEE International Conference on Systems, Man and Cybernetics*, volume 3, pages 6–pp. IEEE, 2002.
- [24] Y. Tan, Y. Shi, F. Buarque, A. Gelbukh, S. Das, and A. Engelbrecht. *Advances in Swarm and Computational Intelligence: 6th International Conference*. Number 2 in Lecture Notes in Computer Science. Springer International Publishing, Beijing, China, 2015.
- [25] M. Ney. Simulation électromagnétique outils de conception. *Techniques Ingénieur*, TIB277DUO(e1030), 2013.
- [26] D.S. Huang, V. Bevilacqua, and P. Premaratne. *Intelligent Computing Theory: 10th International Conference*. Lecture Notes in Computer Science. Springer International Publishing, Taiyuan, China, 2014.
- [27] C. Grosan, A. Abraham, and H. Ishibuchi. *Hybrid Evolutionary Algorithms*. Studies in Computational Intelligence. Springer Berlin, Heidelberg, Germany, 2007.
- [28] M.M. Noel and T.C. Jannett. Simulation of a new hybrid particle swarm optimization algorithm. In *System Theory, 2004. Proceedings of the Thirty-Sixth Southeastern Symposium on System Theory*, pages 150–153. IEEE, 2004.

- [29] A.H. Gandomi and A.H. Alavi. Krill herd: a new bio-inspired optimization algorithm. *Communications in Nonlinear Science and Numerical Simulation*, 17(12):4831–4845, 2012.
- [30] Virtual Library of Simulation Experiments. <https://www.sfu.ca/~ssurjano/ackley.html>. en ligne, consulte le 05/05/2018.
- [31] Virtual Library of Simulation Experiments. <https://www.sfu.ca/~ssurjano/rastr.html>. en ligne, consulte le 05/05/2018.
- [32] E.B. Smith. *Basic Chemical Thermodynamics*. Imperial College Press, London, UK, 2004.
- [33] S.R. Turns. *Thermodynamics: Concepts and Applications*. Number vol. 1 in Thermodynamics: Concepts and Applications. Cambridge University Press, Cambridge, UK, 2006.
- [34] M. Kaufman. *Principles of Thermodynamics*. Undergraduate Chemistry: A Series of Textbooks. CRC Press, Boca Raton, FL, USA, 2002.
- [35] P. Richet. *The Physical Basis of Thermodynamics: With Applications to Chemistry*. Springer US, New York, NY, USA, 2012.
- [36] NASA - Second Law of Thermodynamics. <https://www.grc.nasa.gov/www/k-12/airplane/thermo2.html>. en ligne, consulte le 17/04/2018.
- [37] E. Rathakrishnan. *Fundamentals of Engineering Thermodynamics*. PHI Learning, New Delhi, India, 2005.
- [38] D.R. Gaskell and D.E. Laughlin. *Introduction to the Thermodynamics of Materials, Sixth Edition*. CRC Press, Boca Raton, FL, USA, 2017.
- [39] L. Lue. *Chemical Thermodynamics*. Ventus Publishing, Frederiksberg, Denmark, 2009.
- [40] H. Morowitz. *Foundations of Bioenergetics*. Elsevier Science, London, UK, 2012.
- [41] J.L. Burgot. *The Notion of Activity in Chemistry*. Springer International Publishing, Cham, Switzerland, 2016.
- [42] P. Atkins, J. De Paula, and J. Keeler. *Atkins' Physical Chemistry*. Oxford University Press, Oxford, UK, 2018.
- [43] Vapor Liquid Equilibrium (VLE) - A Guide. <http://web.mit.edu/10.213/www/handouts/vle.pdf>. en ligne, consulte le 25/04/2018.
- [44] K.D. Dahm and D.P. Visco. *Fundamentals of Chemical Engineering Thermodynamics, SI Edition*. Cengage Learning, Stamford, CT, USA, 2014.
- [45] D. Reible. *Fundamentals of Environmental Engineering*. CRC Press, Boca Raton, FL, USA, 2017.
- [46] C.L. Yaws. *The Yaws Handbook of Vapor Pressure: Antoine Coefficients*. Elsevier Science, Oxford, UK, 2015.
- [47] H. Renon and J.M. Prausnitz. Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE journal*, 14(1):135–144, 1968.

- [48] J.M. Prausnitz, R.N. Lichtenthaler, and E.G. de Azevedo. *Molecular Thermodynamics of Fluid-Phase Equilibria*. Pearson Education, NJ, USA, 1998.
- [49] G.M. Kontogeorgis and G.K. Folas. *Thermodynamic Models for Industrial Applications: From Classical and Advanced Mixing Rules to Association Theories*. Wiley, Chichester, UK, 2009.
- [50] E.C. Carlson. Don't gamble with physical properties for simulations. *Chemical engineering progress*, 92(10):35–46, 1996.
- [51] J.P. Wauquier. *Raffinage du pétrole (Le). Tome 2. Procédés de séparation*. Raffinage du pétrole. Editions Technip, Paris, France, 2008.
- [52] J. Gmehling and U. Onken. *Vapor-liquid Equilibrium Data Collection: Aqueous-organic systems*. Chemistry data series. DECHEMA, Deutsche Gesellschaft für Chemisches Apparatewesen, Frankfurt, Germany, 2000.
- [53] R.W. Rousseau. *Handbook of Separation Process Technology*. A Wiley-Interscience publication. Wiley, New York, NY, USA, 1987.
- [54] V.S. Parekh and P.M. Mathias. Efficient flash calculations for chemical process design—extension of the boston–britt “inside–out” flash algorithm to extreme conditions and new flash types. *Computers & chemical engineering*, 22(10):1371–1380, 1998.
- [55] D. Ganderton and J.B. Stenlake. *Unit Processes in Pharmacy: Pharmaceutical Monographs*. Pharmaceutical monographs. Elsevier Science, London, UK, 2014.
- [56] F.B. Petlyuk. *Distillation Theory and its Application to Optimal Design of Separation Units*. Cambridge Series in Chemical Engineering. Cambridge University Press, Cambridge, UK, 2004.
- [57] J. Blacker and M.T. Williams. *Pharmaceutical Process Development: Current Chemical and Engineering Challenges*. Drug Discovery. Royal Society of Chemistry, 2011.
- [58] K. Sambamurthy. *Pharmaceutical Engineering*. New Age International (P) Limited, New Delhi, India, 2007.
- [59] V. Gerbaud. Distillation de mélanges non idéaux. *Techniques Ingénieur*, 2010.
- [60] A. Gorak and E. Sorensen. *Distillation: Fundamentals and Principles*. Handbooks in Separation Science. Elsevier Science, London, UK, 2014.
- [61] J.B. Saint-Hilaire. *Météorologie d’Aristote*. Oeuvres d’Aristote. Ladrance, Paris, France, 1863.
- [62] N.P. Cheremisinoff. *Handbook of Chemical Processing Equipment*. Elsevier Science, Woburn, MA, USA, 2000.
- [63] D. Ebbing and S.D. Gammon. *General Chemistry, Enhanced Edition*. Cengage Learning, Stamford, CT, USA, 2010.
- [64] Cours de rectification: colonnes à plateaux. http://www.azprocede.fr/Cours_GC/rectification_plateaux.html. en ligne, consulte le 19/08/2018.
- [65] U. Diwekar. *Batch Processing: Modeling and Design*. Taylor & Francis, Boca Raton, FL, USA, 2014.

DEMANDE D'IMPRIMATUR

Date de soutenance : 27/09/2018

**DIPLOME D'ETAT DE DOCTEUR
EN PHARMACIE**

présenté par : Kevin FRANIATTE

Sujet :
APPLICATION D'UN BIOALGORITHME POUR LE
DEVELOPPEMENT D'UN PROCEDE DE DISTILLATION
INDUSTRIELLE

Jury :

Président : M. Michel BOISBRUN, Maître de Conférences
Directeur : M. Jean-Etienne FORTIER, Pharmacien-ingénieur
Juges : M. Loïc RONIN, Pharmacien-ingénieur
M. Nicolas BALLOIS, Pharmacien-ingénieur

Vu,

Nancy, le 21 août 2018

Le Président du Jury

Directeur de Thèse

M. Michel BOISBRUN

M. Jean-Etienne FORTIER




Vu et approuvé,

Nancy, le 21.08.2018

Doyen de la Faculté de Pharmacie
de l'Université de Lorraine,



Vu,

Nancy, le

Le Président de l'Université de Lorraine,



N° d'enregistrement : 10385

N° d'identification :

TITRE :

APPLICATION D'UN BIOALGORITHME POUR LE DEVELOPPEMENT D'UN PROCEDE DE DISTILLATION INDUSTRIELLE

Thèse soutenue le 27 Septembre 2018

Par Kevin Franiatte

RESUME :

Ce travail est issu d'une initiative personnelle appliquée à une problématique de développement rencontrée au sein du département de Recherche et Développement d'UCB Pharma à Braine l'Alleud en Belgique. Il consiste à modéliser un procédé de distillation impliqué dans l'étape finale de la synthèse d'un antiépileptique en cours de développement et de proposer des solutions adaptées. Ces modèles utilisés font appel à des notions de thermodynamique décrivant les mécanismes en jeu et aboutissent à des fonctions mathématiques complexes à résoudre. A l'aide d'un algorithme spécialement conçu et basé sur des comportements biologiques, la minimisation de ces fonctions est possible. Il est ainsi possible de générer les outils nécessaires pour étudier différentes stratégies de réduction de la consommation en solvants au cours de cette étape. Des résultats à grande échelle sont également présentés afin de pouvoir discuter de la pertinence des modèles employés.

MOTS CLES : industrie, développement, bioalgorithme, distillation, thermodynamique

Directeur de thèse	Intitulé du laboratoire	Nature
Mr Jean-Etienne FORTIER	UCB Pharma, Braine l'Alleud, Belgique	<input checked="" type="checkbox"/> Expérimentale <input type="checkbox"/> Bibliographique <input checked="" type="checkbox"/> Travail personnel

Thème	<input checked="" type="checkbox"/> Sciences fondamentales <input type="checkbox"/> Hygiène - Environnement <input checked="" type="checkbox"/> Médicament	<input type="checkbox"/> Alimentation - Nutrition <input type="checkbox"/> Biologie <input checked="" type="checkbox"/> Pratique professionnelle
-------	--	--