



HAL
open science

Réflexions autour du profilage physico-chimique des stupéfiants

Laurence Dujourdy

► **To cite this version:**

Laurence Dujourdy. Réflexions autour du profilage physico-chimique des stupéfiants: Ouverture aux domaines de l'agronomie et de l'agroalimentaire. Chimie analytique. Université de Bourgogne (UB), Dijon, FRA.; Ecole Doctorale CARNOT-PASTEUR, 2022. tel-03916287

HAL Id: tel-03916287

<https://hal.science/tel-03916287>

Submitted on 30 Dec 2022

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.

Habilitation à diriger des recherches

Ecole Doctorale CARNOT-PASTEUR
Université de Bourgogne – Franche-Comté
Spécialité CHIMIE THEORIQUE, PHYSIQUE, ANALYTIQUE

Laurence DUJOURDY
L'Institut Agro Dijon
Direction Scientifique - Cellule d'Appui à la Recherche en
Science des données

Réflexions autour du profilage physico- chimique des stupéfiants Ouverture aux domaines de l'agronomie et de l'agroalimentaire

SOUTENUE LE 8 décembre 2022 DEVANT LE JURY COMPOSE
DE :

RAPPORTEURS

Pr. Jean **LOBRY**, Université Claude Bernard Lyon I, LBBE
Pr. Éric **MARCHIONI**, Université de Strasbourg - Institut
Pluridisciplinaire Hubert Curien UMR 7178
Pr. Christophe **CHAMPOD**, Université de Lausanne – École
des Sciences Criminelles

EXAMINATEURS

Pr. Evelyne **VIGNEAU**, Oniris- École Nationale Vétérinaire,
Agro-alimentaire et de l'Alimentation Nantes Atlantique -
StatSC (unité de Statistique, Sensométrie et Chimiométrie)
Dr. Anne-Béatrice **DUFOUR**, Université Claude Bernard Lyon
I, LBBE
Pr. Pierre-Yves **LOUIS**, Université de Bourgogne, UMR PAM



Remerciements

Je suis très reconnaissante aux membres du jury qui m'ont fait l'honneur de juger ce travail.

Tout au long de mon parcours professionnel, j'ai eu la chance d'être toujours soutenue et encouragée par mes parents et mon frère ainsi que par mes proches amis et ma famille. Ils ont toujours été présents pour m'épauler lors des moments de doute et me redonner de l'énergie lors des baisses de régime. Un grand merci pour vos paroles réconfortantes ou stimulantes, vos attentions, pour le partage de bons « p'tits » repas, de moments exquis, de virées, de fous rires ... malgré la distance et les années.

Je tiens également à remercier mes relecteurs de la 1^{re} heure dont la tâche était ardue, car le document était brut et plutôt dense.

Je remercie aussi tous les collègues de tous horizons, rencontrés à travers les missions, les réunions, anciens ou récents, de tous les jours ou de temps en temps, pour les riches discussions qui m'ont permis de voir les choses autrement et de me faire avancer. Ce mémoire me fournit aussi l'occasion d'exprimer de sincères remerciements à tous les collaborateurs qui de près ou de loin ont participé à ces travaux. J'ai une pensée émue pour ceux disparus trop tôt.

Préambule

Ce manuscrit, rédigé en vue de l'obtention de l'Habilitation à Diriger des Recherches, présente une synthèse de mes travaux depuis mon Post-Doctorat (2000-2003) à l'université de Lausanne jusqu'à ce jour (juillet 2022).

Il comporte trois chapitres.

- Le premier est un curriculum vitae qui résume mon parcours professionnel, mes activités scientifiques et pédagogiques ;
- Le second chapitre détaille mon travail de recherche passé
- et le troisième chapitre donne l'orientation de mes futurs travaux.

Table des matières

Chapitre 1 : Curriculum Vitae	Erreur ! Signet non défini.	10
Chapitre 2 : Expériences en matière de recherche		18
2.1. Introduction		18
2.2. Quelques définitions et éléments de contexte		18
2.2.1. Que signifie être chercheur dans un cadre pénal ?		18
2.2.2. Qu'entend-on par profilage physico-chimique ?		20
2.2.3. Qu'est-ce la chimiométrie ?		21
2.2.4. Qu'est-ce qu'une donnée ?		22
2.3 : Caractérisation des stupéfiants et profilage		23
2.3.1 Éléments de contexte		23
a) Qu'est-ce qu'une drogue ?		23
b) Classification des drogues par leurs effets		24
c) Contexte juridique dans lequel sont analysées les saisines		25
2.3.2. Objectifs des analyses		27
a) Étape 1 : l'identification		28
b) Étape 2 : le dosage (la quantification)		31
c) Étape 3 : les comparaisons		32
2.4. Échantillonnage de stupéfiants		34
2.4.1. Échantillonnage pour des analyses d'identification (qualitatives)		34
2.4.2. Échantillonnage pour des analyses quantitatives (dosages)		40
a) Première phase d'étude : l'hétérogénéité des drogues		43
b) Deuxième phase d'étude : effet de la taille des particules		45
c) Troisième phase d'étude : plan d'échantillonnage général		46
2.5. Validation de méthodes analytiques		48
2.5.1. Cadre de la norme ISO 17025		49
2.5.2. Principe du profil d'exactitude		51
a) Base des calculs statistiques mis en œuvre pour la prise de décision		52
b) Mise en place de l'application VALMETH		57
2.6. Profilage chimique des stupéfiants		61
2.6.1. Processus de profilage		62
2.6.2. Genèse du profilage chimique des drogues		63

2.6.3. Principe de la comparaison de profils chimiques	64
2.6.3. Évaluation des mesures de dissimilarités	69
2.7. Conclusion	113
Chapitre 3 : Réflexions sur les activités de recherche et perspectives	118
3.1.Changement de cap, des stupéfiants à l'agronomie et à l'agroalimentaire	115
3.2 Perspectives à court et moyen terme	118
3.2.1 À court terme.....	118
3.2. A moyen terme.....	120
CONCLUSION.....	127
Références bibliographiques	128
ANNEXES	136

Table des figures

Figure 1 : Procédé global d'analyses	21
Figure 2 : Lien entre la chimiométrie et d'autres disciplines.....	22
Figure 3 : Catégories des techniques analytiques rencontrées dans l'analyse de stupéfiants	29
Figure 4 : Processus d'identification d'une substance supposée illicite	29
Figure 5 : Évolution de la forme chimique au cours du temps	30
Figure 6 : Étape de dosage	31
Figure 7 : Concept global du procédé de profilage des drogues.....	33
Figure 8 : Chronologie de la réflexion sur l'échantillonnage quantitatif et production scientifique.	41
Figure 9 : Aperçu du processus d'échantillonnage et d'analyse quantitative.	43
Figure 10 : Illustration de l'hétérogénéité de constitution d'un échantillon d'herbe de cannabis.....	44
Figure 11 : Illustration de l'hétérogénéité de distribution avec l'échantillon d'herbe de cannabis	44
Figure 12 : Comparaison entre un échantillonnage arbitraire et un échantillonnage incrémental.	45
Figure 13 : Carte de contrôle sur l'étendue, carte R.	48
Figure 14 : Échantillonnage = phase préliminaire à l'analyse.	48
Figure 15 : Courbe d'étalonnage permettant d'obtenir la valeur de la concentration d'un analyte	50
Figure 16 : Illustration d'un point de vue statistique de l'objectif de la validation.	51
Figure 17 : Illustration de la différence de prise de décision suivant les deux approches statistiques utilisées.	54
Figure 18 : Illustration qu'une bonne méthode ne donne pas forcément de bons résultats.	55
Figure 19 : Illustration du profil d'exactitude exprimé par le taux de recouvrement.	56
Figure 20 : Page d'accueil de l'application interne VALMETH.....	58
Figure 21 : Exemple de résultats d'analyses pour la validation de la méthode de dosage de l'héroïne.	59
Figure 22 : Exemple d'indices calculés pour l'aide à la décision.	60
Figure 23 : Profil d'exactitude et profil de risque pour le dosage de l'héroïne	60
Figure 24 : Processus de science forensique dans le cas du profilage de drogues.....	61
Figure 25 : Illustration du processus général de profilage.....	62
Figure 26 : Principe de comparaison de 2 profils chimiques	63
Figure 27 : Méthodes d'analyses multivariées.....	65
Figure 28 : Différences entre les distances.....	66
Figure 29 : Effet de la dilution des drogues sur la forme du chromatogramme.	68
Figure 30 : Visualisation des distributions des distances des 2 ensembles d'échantillons.	69
Figure 31 : Représentation du calcul des faux positifs et des faux négatifs.....	70
Figure 32 : Voies de synthèse communes de l'amphétamine.	75
Figure 33 : Synthèse de la MDMA à partir du safrole	80
Figure 34 : Exemples de logo de comprimés d'ecstasy.....	81
Figure 35 : ACP d'échantillons de rue provenant de différents pays	82
Figure 36 : Voies de synthèse	85
Figure 37 : Copie d'écran de la feuille de démarrage de la macro	86
Figure 38 : Exemple de sortie sous forme de matrice	87
Figure 39 : ACP d'échantillons de MA	88
Figure 40 : Dendrogramme de la CAH.....	89
Figure 41 : Pavot somnifère ou pavot à opium.....	91
Figure 42 : Étapes de fabrication de l'héroïne à partir du pavot (<i>Papaver somniferum</i> L.).....	92
Figure 43 : Cocaïer <i>Erythroxylum coca</i> Lam.....	95
Figure 44 : Étapes de fabrication de la cocaïne à partir de la feuille de coca.	96
Figure 45 : Chromatogrammes d'un échantillon d'héroïne analysé sur 2 appareils différents	99
Figure 46 : Distribution et boîtes à moustache des aires de la noscapine pour les 2 appareils.....	99
Figure 47 : Mouvement et distribution des drogues illicites du fabricant au consommateur	101
Figure 48 : Principe et visualisation des liens chimiques avec i2 Analyst Notebook®	102
Figure 49 : Exemple d'un chromatogramme d'un échantillon de cocaïne analysé par HS-GC-FID.....	104
Figure 50 : Graphiques des individus et des variables de la PLS-DA	105
Figure 51 : Exemple fictif du calcul du rapport de vraisemblance.....	106
Figure 52 : Illustration des liens chimiques	110
Figure 53 : Principales routes du trafic de la cocaïne, source UNODC.	111
Figure 54 : Flux du trafic d'héroïne dans une affaire en particulier (Strasbourg).	112
Figure 55 : Trois étapes des analyses.....	113
Figure 56 : Analyses en composantes principales des profils chimiques des solvants de la cocaïne	124
Figure 57 : Visualisation t-SNE. Graphique en 2 dimensions [A] et en 3D [B],.....	125

Chapitre 1

Curriculum vitae

Cette partie a été volontairement ôtée



LISTE DES TRAVAUX PUBLIÉS

55 publications - h index = 17 (2022) cf. détail ci-après

ORCID n° : 0000-0001-8318-6979

Web of Science Researcher ID AAC-8384-2022

Reviewer dans les revues internationales depuis 2006 : Forensic Science International (108 articles révisés), Analytical & Bioanalytical Chemistry (7 articles), Drug testing and Analysis (2 articles), Analytical Chemistry (1 article), Journal of Chromatography A (1 article), Science & Justice (1 article), Molecules (2 articles)

Caractéristiques des journaux dans lesquels les articles ont été principalement publiés

Drug Testing and Analysis : impact factor 2.80 ; SCImago Journal Rank (SJR) 0.75 ; H-index : 54 ; Subject Area: Analytical Chemistry (Q1); Spectroscopy (Q1); Pharmaceutical Science (Q1) ; Environmental chemistry (Q2)

Food Chemistry : impact factor 5.68 ; SCImago Journal Rank (SJR) 1.768 ; H-index : 262 ; Subject Area: Analytical Chemistry (Q1); Food Science (Q1); Medicine (miscellaneous) (Q1)

Forensic Science International : impact Factor: 2.22 ; SCImago Journal Rank (SJR) 0.879; H-index : 120 ; Subject Area: Pathology and Forensic Medicine (Q1)

Frontiers in Microbiology: impact factor 4.19 ; SCImago Journal Rank (SJR) 1.633 ; H-index : 135 ; Subject Area: Microbiology (Q1); Microbiology (medical) (Q1)

Journal of Chromatography A : impact factor 3.81 ; SCImago Journal Rank (SJR) 1.188 ; H-index : 229 ; Subject Area: Analytical Chemistry (Q1); Medicine (miscellaneous) (Q1); Organic Chemistry (Q1); Biochemistry (Q2)

Methods and Applications in Fluorescence : impact factor 2.84; ; SCImago Journal Rank (SJR) 0.877 ; H-index : 25 ; Subject Area: Atomic and Molecular Physics, and Optics (Q1); Instrumentation (Q1); Materials Science (miscellaneous) (Q1); Spectroscopy (Q1)

Polymer Bulletin : impact factor 1.66 ; SCImago Journal Rank (SJR) 0.40 ; H-index : 60 ; Subject Area: Chemistry miscellaneous) (Q2); Materials Chemistry (Q2); Polymers and Plastics (Q2); Condensed Matter Physics (Q3)

Polymer International : impact factor 2.50 ; SCImago Journal Rank (SJR) 0.603 ; H-index : 105 ; Subject Area: Polymers and Plastics (Q1); Materials Chemistry (Q2); Organic Chemistry (Q2)

Science & Justice : impact factor 2.07 ; SCImago Journal Rank (SJR) 0.67 ; H-index : 42 ; Subject Area: Pathology and Forensic Medicine (Q2)

Sensors : impact factor 3.51 ; SCImago Journal Rank (SJR) 0.592 ; H-index : 172 ; Subject Area: Analytical Chemistry (Q2); Atomic and Molecular Physics, and Optics (Q2); Electrical and Electronic Engineering (Q2); Instrumentation (Q2); Medicine (miscellaneous) (Q2); Biochemistry (Q3)

Talanta : impact factor 4.93 ; SCImago Journal Rank (SJR) 1.152; H-index : 161 ; Subject Area: Analytical Chemistry (Q1); Spectroscopy (Q1); Biochemistry (Q2)

Thématique Profilage chimique : statistiques appliquées / chimiométrie

Revue à comité de lecture (26 articles)

- (SA1) J. Melendrez-Ruiz, **L. Dujourdy**, I. Goisbault, J-C. Charrier, K. Pagnat, S. Nicklaus, G. Arvisenet, S. Chambaron, "You look at it, but will you choose it": Is there a link between the foods consumers look at and what they ultimately choose in a virtual supermarket? **Food Quality and Preference**, 98, pp.104510, 2022, <https://doi.org/10.1016/j.foodqual.2021.104510>
- (SA2) S. Pagès-Hélary, **L. Dujourdy**, N., Cayot, "Flavor compounds in Blackcurrant berries: multivariate analysis of datasets obtained with natural variability and various experimental parameters", **LWT**, 153, 2022, <https://doi.org/10.1016/j.lwt.2021.112425>
- (SA3) S. Pagès-Hélary, **L. Dujourdy**, N., Cayot, "Identification of Volatile Compounds in Blackcurrant Berries: Differences among Cultivars ", **Molécules**, 26(20):6254, 2021, <https://doi.org/10.3390/molecules26206254>
- (SA4) J. Melendrez-Ruiz, I. Goisbault, J-C. Charrier, K. Pagnat, **L. Dujourdy**, G. Arvisenet, S. Chambaron, "An Exploratory Study Combining Eye-Tracking and Virtual Reality: Are Pulses Good "Eye-Catchers" in Virtual Supermarket Shelves?", **Front. Virtual Real.** 2:655273. 2021, <https://doi.org/10.3389/frvir.2021.655273>
- (SA5) N. Lebbos, C. Keller, **L. Dujourdy**, M. Afram, P. Curmi, T. Darwish, C. Daou, E. Bou-Maroun, "Validation of a new method for monitoring trace elements in Mediterranean cereal soils", **International Journal of Environmental Analytical Chemistry**, 2021, <https://doi.org/10.1080/03067319.2021.1953000>
- (SA6) G. Duley, **L. Dujourdy**, S. Klein, A. Werwein, C. Spartz, R. D Gougeon, D. K. Taylor, "Regionality in Australian Pinot Noir wines: a study on the use of NMR, ICP-MS, and Spectroscopy on commercial wines", **Food Chemistry**, 340, 2021, <https://doi.org/10.1016/j.foodchem.2020.127906>
- (SA7) J. Rossignol, **L. Dujourdy**, D. Stuerga, P. Cayot, R. Gougeon, E. Bou-Maroun, "A First Tentative for Simultaneous Detection of Fungicides in Model and Real Wines by Microwave Sensor Coupled to Molecularly Imprinted Sol-Gel Polymers", **Sensors**, 20(21), 6224, 2020, <https://doi.org/10.3390/s20216224>
- (SA8) A. Anchieta Câmara, T. Dat Nguyen, R. Saurel, C. Sandt, C. Peltier, **L. Dujourdy**, F. Husson, "Biophysical Stress Responses of the Yeast *Lachancea thermotolerans* During Dehydration Using Synchrotron-FTIR Microspectroscopy", **Frontiers in Microbiology**, Vol 11, pp.899-913, 2020
- (SA9) C. Modugno, C. Peltier, H. Simonin, **L. Dujourdy**, F. Capitani, C. Sandt, J-M. Perrier-Cornet, "Understanding the effects of high pressure on bacterial spores using synchrotron infrared spectroscopy", **Frontiers in Microbiology**, Vol 10, pp. 3122-3132, 2020
- (SA10) N. Lebbos, C. Daou, R. Ouaini, H. Chebib, M. Afram, P. Curmi, **L. Dujourdy**, E. Bou-Maroun, M-C. Chagnon, "Lebanese Population Exposure to Trace Elements via White Bread Consumption", **Foods**, 8(11), pp. 574-594, 2019
- (SA11) S. Daoud, E. Bou-Maroun, **L. Dujourdy**, G. Waschatko, N. Billecke, P. Cayot, "Fast and direct analysis of oxidation levels of oil-in-water emulsions using ATR-FTIR", **Food Chemistry**, 293, pp. 307-314, 2019
- (SA12) C. Peltier, P. Winckler, **L. Dujourdy**, S. Bechoua, JM. Perrier-Cornet. "Analysis of multivariate images in fluorescence microscopy", **Methods and Applications in Fluorescence**, IOP Science 7 (3), pp. 035004, 2019

- (SA13) **L. Dujourdy**, C. Charvoz, M. Dalmasso, A-B. Dufour, "What a validation strategy means for the quantitation of cocaine and heroin?", **Forensic Science International**, 251, p. 32-39, 2015
- (SA14) T. Csesztregi, M. Bovens, **L. Dujourdy**, A. Franc, J. Nagy, "Sampling of illicit drugs for quantitative analysis – Part III: Sampling plans and sample preparations", **Forensic Science International**, 241, p. 212-219, 2014
- (SA15) M. Bovens, T. Csesztregi, A. Franc, J. Nagy, **L. Dujourdy**, "Sampling of illicit drugs for quantitative analysis. Part II: Study of particle size and its influence on mass reduction", **Forensic Science International**, 234, p. 174–180, 2014
- (SA16) **L. Dujourdy**, C. Charvoz, M. Dalmasso, A-B. Dufour, "A la recherche du juste dosage : statistique appliquée à l'analyse de la cocaïne", **Annales Pharmaceutiques Françaises**, 71(3), p. 193–200, 2013
- (SA17) **L. Dujourdy**, T. Csesztregi, M. Bovens, A. Franc, J. Nagy, "Sampling of Illicit Drugs for Quantitative Analysis – Part I. Heterogeneity Study of Illicit Drugs in Europe", **Forensic Science International**, 231(1–3), p. 249–256, 2013
- (SA18) Bolck, C. Weyermann, **L. Dujourdy**, P. Esseiva, J. van den Berg, "Different likelihood ratio approaches to evaluate the strength of evidence of MDMA tablets comparison", **Forensic Science International**, 191(1-3), p. 42-51, 2009
- (SA19) **L. Dujourdy**, F. Besacier, "Headspace profiling of cocaine samples for intelligence purposes", **Forensic Science International**, 179(2-3), p. 111-122, 2008
- (SA20) **L. Dujourdy**, V. Dufey, F. Besacier, N. Miano, R. Marquis, E. Lock, L. Aalberg, S. Dieckmann, F. Zrcek, J.S. Bosenko Jr., "Drug intelligence based on organic impurities in illicit MA samples", **Forensic Science International**, 177(2-3), p. 153-161, 2008
- (SA21) C. Weyermann, C. Delaporte, P. Esseiva, R. Marquis, **L. Dujourdy**, E. Lock, L. Aalberg, S. Dieckmann, F. Zrcek, J.S. Bosenko Jr., "Drug intelligence based on organic impurities in MDMA tablets", **Forensic Science International**, 177(1), pp. 11-16, 2008
- (SA22) V. Dufey, **L. Dujourdy**, F. Besacier, H. Chaudron, "A quick and automated method for profiling heroin samples for tactical intelligence purposes", **Forensic Science International**, 169(2-3), pp. 108-117, 2007
- (SA23) K. Andersson, E. Lock, K. Jalava, H. Huizer, S. Jonson, E. Kaa, A. Lopes, A. Poortman-van der Meer, E. Sippola, **L. Dujourdy**, J. Dahlén, "Development of a harmonised method for the profiling of amphetamines VI: Evaluation of methods for comparison of amphetamine", **Forensic Science International**, 169(1), p. 86-99, 2007
- (SA24) P. Esseiva, F. Anglada, **L. Dujourdy**, F. Taroni, P. Margot, E. Du Pasquier, M. Dawson, C. Roux, P. Doble, "Chemical profiling and classification of illicit heroin by principal component analysis, calculation of inter sample correlation and artificial neural networks", **Talanta** 67 (2), p. 360-367, 2005
- (SA25) P. Esseiva, **L. Dujourdy**, F. Anglada, F. Taroni, P. Margot, "A methodology for illicit heroin seizures comparison in a drug intelligence perspective using large databases", **Forensic Science International**, 132 (2), p. 139-152, 2003
- (SA26) **L. Dujourdy**, Giulia Barbati, F. Taroni, P. Esseiva, F. Anglada, O. Guéniat, P. Margot, "Evaluation of links in heroin seizures", **Forensic Science International**, 131 (2-3), p. 171-183, 2003

Reuves sans comité de lecture (2 articles)

- **L. Dujourdy**, A. Schvartz, B. Larbre, Y. Le Fur, "L'évaluation sensorielle au service d'une nouvelle offre de verres à vin", **Territoires du vin** [En ligne], 10 | 2019, mis en ligne le 04 juillet 2019, URL : <http://preo.u-bourgogne.fr/territoiresduvin/index.php?id=1724>
- **L. Dujourdy**, F. Besacier, T. Soto, "La police scientifique fait des statistiques", **Annales des Falsifications, de l'Expertise Chimique & Toxicologique**, 974, p. 39-49, 2011

Thématique Profilage chimique : chimie analytique - analyse de stupéfiants

Reuves à comité de lecture (10 articles)

- (CA1) Onoka, A. Banyikwa, P. N Banerjee, J. J Makangara, **L. Dujourdy**, "A Review of the Newly Identified Impurities Profiles in Methamphetamine Seizures", **Forensic Science International Synergy**, Vol. 2, pp. 194-205, 2020

- (CA2) V. Ladroue, **L. Dujourdy**, F. Besacier, P. Jame, "IRMS to study a common cocaine cutting agent: phenacetin", **Drug Testing and Analysis**, 9 (3), pp. 479-484, 2017.
- (CA3) **L. Dujourdy**, F. Besacier, "A study of cannabis potency in France over a 25 years period (1992–2016)", **Forensic Science International**, 272, pp. 72-80, 2017
- (CA4) M. Marchini, C. Charvoz, **L. Dujourdy**, N. Baldovini, J-J. Filippi, "Multidimensional analysis of cannabis volatile constituents: Identification of 5,5-dimethyl-1-vinylbicyclo[2.1.1]hexane as a volatile marker of hashish, the resin of Cannabis sativa L", **Journal of Chromatography A**, Volume 1370, p. 200-215, 2014
- (CA5) **L. Dujourdy**, F. Besacier, "L'héroïne saisie en France. Données statistiques issues de la base nationale des Laboratoires de Police Scientifique", **Annales Pharmaceutiques Françaises**, 68 (2), p. 127-132, 2010
- (CA6) S. Locicero, P. Esseiva, P. Hayoz, **L. Dujourdy**, F. Besacier, P. Margot, "Cocaine profiling for strategic intelligence purposes, a cross-border project between France and Switzerland: Part II. Validation of the statistical methodology for the profiling of cocaine", **Forensic Science International**, 177, p. 199-206, 2008
- (CA7) S. Locicero, P. Hayoz, P. Esseiva, **L. Dujourdy**, F. Besacier, P. Margot, "Cocaine profiling for strategic intelligence purposes, a cross-border project between France and Switzerland: Part I. Optimisation and harmonisation of the profiling method", **Forensic Science International**, 167(2-3), p. 220-228, 2007
- (CA8) P. Mura, B. Brunet, **L. Dujourdy**, C. Paetzold, G. Bertrand, B. Sera, Benoît Saclier, M. Deveaux, G. Pépin, M. Perrin, Y. Lecompte, V. Dumestre Toulet, V. Cirimele, P. Kintz, "Cannabis d'hier et cannabis d'aujourd'hui. Augmentation des teneurs en THC de 1993 à 2004 en France", **Annales de Toxicologie Analytiques**, 1 (XVIII), p. 3-6, 2006
- (CA9) P. Gimeno, F. Besacier, M. Bottex, **L. Dujourdy**, H. Chaudron-Thozet, "A study of impurities in intermediates and 3,4-methylenedioxymethamphetamine (MDMA) samples produced via reductive amination routes", **Forensic Science International**, 155 (2-3), p. 141-157, 2005
- (CA10) J. Ballany, B. Caddy, M. Cole, Y. Finnon, L. Aalberg, K. Janhunen, E. Sippola, K. Andersson, C. Bertler, J. Dahlén, I. Kopp, **L. Dujourdy**, E. Lock, P. Margot, H. Huizer, A. Poortman, E. Kaa, A. Lopes. "Development of a harmonised pan-European method for the profiling of amphetamines", **Science & Justice**, Volume 41, Issue 3, 2001

Reuves sans comité de lecture (10)

- **L. Dujourdy**, F. Besacier, C. Charvoz, V. Ladroue, « L'Institut National de Police Scientifique, la contribution d'un service répressif sur l'axe prévention », **La Lettre du Réseau de Prévention des Addictions** (Respad), n°22, 2015
- **L. Dujourdy**, B. Santin, T. Soto, "Chimie des couleurs et investigations préliminaires en police scientifique", **L'Actualité Chimique**, N°378-379, p. 52-61, octobre-novembre 2013
- **L. Dujourdy**, T. Soto, " Cannabis « bio » ou cannabinoïdes de synthèse ? ", **L'Actualité Chimique**, N°378-379, p. 106-111, octobre-novembre 2013
- **L. Dujourdy**, F. Besacier, V. Ladroue, "La cocaïne saisie en France. Données statistiques issues de la base nationale de l'Institut National de Police Scientifique", **L'Actualité Chimique**, N° 342-343 p. 29-37, juin-juillet-août 2010
- V. Ladroue, **L. Dujourdy**, F. Besacier, "Profilage des impuretés de l'amphétamine", **L'Actualité Chimique**, , N° 342-343 juin-juillet-août, p. 37-45, 2010
- F. Besacier, **L. Dujourdy**, T. Soto, "Outils et méthodologies de police scientifique pour l'étude des drogues", **Cahiers de la Sécurité**, n°5, p. 81-87, 2008
- S. Locicero, **L. Dujourdy**, P. Esseiva, F. Besacier, P. Hayoz, F. Anglada, P. Margot et J. Guitton. "Apport scientifique à la lutte contre le phénomène transfrontalier de stupéfiants. Mise en réseau de l'information." **Revue Internationale de Criminologie et de Police Technique et Scientifique**. Volume LVIII N°1, janvier-mars, p. 83-92, 2005
- P. Margot, **L. Dujourdy**, L. Lanzi, F. Botti, R. Marquis, S. Locicero, "Notes de police scientifique". **Revue Internationale de Criminologie et de Police Scientifique**. Volume LVI N°1, janvier-mars p. 117-121, 2003

- F. Anglada, **L. Dujourdy**, I. Baer, D. Alvarez, C. Zingg, C. Brignoli, E. Lock, P. Esseiva, P. Margot, "Les dérivés amphétaminiques en Suisse et dans le monde", **Revue Internationale de Criminologie et de Police Technique et Scientifique**. Volume LV N°4, octobre-décembre, p. 465-490, 2002
- O. Guéniat, F. Anglada, F. Taroni, P. Margot, **L. Dujourdy**, P. Esseiva. "Les signatures chimiques à partir des saisies de produits stupéfiants. Utilité et gestion de l'information", **Revue internationale de criminologie et de police technique et scientifique**. Volume LV N°1, janvier-mars, p. 104-111, 2002

Thématique : Physico-chimie des matériaux polymères et composites

Revue à comité de lecture (4 articles)

- (PC1) S. Locicero, **L. Dujourdy**, W. Mazzella, P. Margot, E. Lock, "Dynamic of the ageing of ballpoint pen inks: quantification of phenoxyethanol by GC-MS", **Science & Justice**, 44(3), p. 165-171, 2004
- (PC2) **L. Dujourdy**, J.P. Bazile, J.P. Cohen Addad, "Crystallinity and amorphous segmental properties in random Propylene-Ethylene copolymers. NMR correlation", **Polymer International**, 48, p. 558-564, 1999
- (PC3) J.P. Cohen Addad, **L. Dujourdy**, "Bimodal Polydimethylsiloxane adsorption on silica aggregates. Characteristic properties", **Polymer Bulletin**, 41, p. 253-260, 1998
- (PC4) **L. Dujourdy**, J.P. Cohen Addad, "Kinetics of polydimethylsiloxane adsorption on silica aggregates. Effect of chain molecular weight and silica concentration", **Eurofillers 95 Extended Abstracts**, p. 259-262, 1995

Ouvrage, chapitres d'ouvrage et rapports scientifiques (5 publications)

Ouvrage

- **L. Dujourdy**, M. Bovens, J. Nagy, T. Csesztregi, Franc A., "Guidelines on sampling of illicit drugs for quantitative analysis", Edition: Issue No: 001, Publisher: **European Network for Forensic Science Institutes (ENFSI)** - Drugs Working Group, April 2014

Chapitres d'ouvrage

- C. Gée, V. Mignon, **L. Dujourdy**, E. Denimal. Chapter 26. "High throughput field phenotyping (HTFP) of wheat and weed cover in field experiments using RGB images: assessment of crop-weed competition with a simple ecophysiological mode" **Precision agriculture '21**, John V. Stafford, pp.225 - 232, 2021
- **L. Dujourdy** dans, "La science forensique - Le futur d'une discipline", **Presses Polytechniques et Universitaires Romandes (PPUR) Collection Sciences forensiques**, 2015, 198 pages

Rapports scientifiques

- L. Ahlberg, K. Andersson, J. Ballany, C. Bertler, M.D. Cole, **L. Dujourdy**, Y. Finnon, E. Lock, K. Jalava, H. Huizerd, S. Jonson, E. Kaa, T. Kärkkäinen, A. Lopes, P. Margot, A. Poortman-van der Meer, E. Sippola, S. Stevenson, P. White, J. Dahlén, "Development of a harmonised method for the profiling of amphetamines." Project SMT – CT98 – 2277, **European Commission**, Final report, ISBN 951-53-2528-5, 2003, Finland.
- **L. Dujourdy**, F. Taroni, P. Margot, "Classification, interpretation and statistical evaluation of data in illicit drug profiling", Rapport scientifique, **Fond National Suisse de la Recherche Scientifique**, n° de requête 1115-068017, 2003.

Rapports scientifiques industriels (confidentiels)

Manufacture des Pneumatiques Michelin

- « Recherche des molécules organiques adsorbées sur les noirs de carbone. Détection par l'observation des protons », **L. Dujourdy**, J. P. Cohen Addad, J. P. Bazile, novembre 96
- « Étude du signal de relaxation des noirs de carbone aux temps courts », **L. Dujourdy**, J. P. Cohen Addad, J. P. Bazile, avril 97

Rhône-Poulenc

- « Mise en évidence par RMN de l'adsorption dans les mélanges PDMS/silice non extraits », **L. Dujourdy**, J. P. Cohen Addad, J. P. Bazile, Bordeaux D., avril 97

Schneider Electric

- « Étude en rhéologie du stockage des mélanges Résine époxyde/Durcisseur/Alumine », **L. Dujourdy**, J. P. Cohen Addad, J. P. Bazile, avril 97

Groupement de Recherches de Lacq – Elf Atochem

- « Étude par RMN du comportement des copolymères statistiques éthylène/propylène granulés en température », **L. Dujourdy**, J. P. Cohen Addad, J. P. Bazile, mai 97
- « Évolution des solutions de mélamine/formol. Approche préliminaires », **L. Dujourdy**, J. P. Cohen Addad, J. P. Bazile, juillet 97
- « Étude par RMN du comportement des copolymères éthylène/propylène », **L. Dujourdy**, J. P. Cohen Addad, J. P. Bazile, septembre 97
- « Étude RMN des copolymères éthylène/propylène, série ZPR et série EC », **L. Dujourdy**, J. P. Cohen Addad, J. P. Bazile, février 98
- « Évolution des solutions de mélamine/formol. Étude par RMN », **L. Dujourdy**, J. P. Cohen Addad, J. P. Bazile, mars 98
- « Rapport final : évolution des solutions de mélamine/formol. Étude par RMN », **L. Dujourdy**, J. P. Cohen Addad, J. P. Bazile, juillet 98

VALORISATION SCIENTIFIQUE : CONGRÈS / COLLOQUES / ATELIERS

46 communications

- **Internationales**

- (I1) European Conference on Precision Agriculture (ECPA), C. Gée, V. Mignon, **L. Dujourdy**, E. Denimal, présentation orale: "High throughput field phenotyping (HTFP) of wheat and weed cover in field experiments using RGB images: assessment of crop-weed competition with a simple ecophysiological model", Budapest, 2021
- (I2) 14th Pangborn Sensory Science Symposium, J. Melendrez-Ruiz, **L. Dujourdy**, I. Goisbault, J-C. Charrier, K. Pagnat, S. Nicklaus, G. Arvisenet, S. Chambaron, présentation orale: "You look at it, but will you choose it?: Is there a link between what consumers look at and what they ultimately choose in a virtual supermarket?", fully online virtual event 2021
- (I3) 9th european conference on sensory and consumer research (eurosense), J. Melendrez-Ruiz, I. Goisbault, J-C. Charrier, K. Pagnat, **L. Dujourdy**, G. Arvisenet, S. Chambaron, poster: "Evaluating nudging strategies towards pulses: how a virtual setting combining eye-tracking and explicit measures can help to understand consumer behavior?", Rotterdam, 2020
- (I4) 9th european conference on sensory and consumer research (eurosense), J. Melendrez-Ruiz, I. Goisbault, J-C. Charrier, K. Pagnat, **L. Dujourdy**, G. Arvisenet, S. Chambaron, poster: "Understanding consumer behavior and evaluating nudge strategies to increase food choice of pulses: how virtual reality combined with eye-tracking and explicit measures can help?", Rotterdam, 2020
- (I5) UNESCO Chair "Culture and Wine Traditions" of the University of Burgundy and the University of Crete, International conference "Vigne, vin, alimentation, santé", **L. Dujourdy**, Y. Le Fur, présentation orale: "Sensory evaluation for a new offer of wine glass", Héraklion, 2018
- (I6) IAFP'S European Symposium on Food Safety, C. Modugno, H. Simonin, C. Peltier, **L. Dujourdy**, JM. Perrier-Cornet, poster: "Reversible initiation of spore germination by high-pressure at 20°C", Stockholm, 2018
- (I7) 8th European Academy of Forensic Science Conference, **L. Dujourdy**, F. Besacier, atelier: "Workshop on data treatment of drug profiles", Lyon, 2018
- (I8) 7th European Academy of Forensic Science Conference, T. Csesztregi, M. Bovens, **L. Dujourdy**, atelier: "Workshop on sampling of illicit drugs for quantitative analysis", Prague, 2015
- (I9) 21st Annual Meeting, ENFSI Working Group Drugs, T. Csesztregi, M. Bovens, **L. Dujourdy**, atelier: "Workshop on sampling of illicit drugs for quantitative analysis", Dublin, 2015
- (I10) 4th International Conference on Chemical Monitoring Programme and Crystal Methamphetamine, **L. Dujourdy**, présentation orale: "Forensic assistance to the fight against precursors trafficking and synthetic drugs production", Cape Town 2013
- (I11) Forensic Drug Expert Meeting, **L. Dujourdy**, présentation orale: "What can be deduced from forensic purity data?", European Monitoring Centre for Drugs and Drug Addiction (EMCCDA), Lisbonne, 2013
- (I12) 17th International Forensic Science Symposium, 2 posters: "EMPHASIS project", "DOGGIES project", INTERPOL, Lyon, 2013
- (I13) 19th annual meeting of ENFSI Working Group Drugs, **L. Dujourdy**, Tamás Csesztregi, Michael Bovens, Anne Franc, Júlia Nagy, présentation orale: "Quantitative Sampling for Illicit Drugs", Dubrovnik 2013
- (I14) Colloque sur le trafic de drogues de synthèse en Asie du Sud-Est, **L. Dujourdy**, présentation orale: "Analysis of synthetic drugs by French forensic laboratories", Hanoï, 2012.
- (I15) 12th annual meeting of Fingerprint Working Group, poster: "CARTES: Expedited and Easy Capture of Latent Prints from crime scene", Lyon 2012
- (I16) 6th European Academy of Forensic Science (EAFS) triennial Meeting, poster: "EMPHASIS: A novel system for pinpointing production of home-made explosives and bombs", The Hague 2012
- (I17) 18th annual meeting of ENFSI Working Group Drugs, F. Besacier, **L. Dujourdy**, présentation orale: "Cannabis profiling", Paphos 2012.

Chapitre 1 : Curriculum vitae

- (I18) 16th International Forensic Science Symposium, 2 posters : "French national narcotics database (S.T.U.P.S.®): update of the computerized application and project management", "Automatic tool for illicit drugs sample comparisons", INTERPOL, Lyon, 2010
- (I19) 60th annual scientific meeting of American Academy of Forensic Sciences (AAFS), F. Bévalot, **L. Dujourdy**, L. Fanton, N. Cartiser, L. Magné, F. Besacier, C. Le Meur, Y. Gaillard, D. Malicier, présentation orale : "Statistic interpretation of meprobamate concentrations in bone marrow, vitreous and bile", Washington 2008
- (I20) 14th meeting of ENFSI Working Group Drugs, F. Besacier, **L. Dujourdy**, présentation orale : "Heroin and Cocaine in France – Trends through 20 years", Paris 2008
- (I21) 4th European Academy of Forensic Science Conference, 3 posters : "Cocaine profiling: Analysis of solvents and alkaloids" ; "Study of the Influence of Cutting Agents on Cocaine Profiling including Tests with Different Derivative Agents" ; "CHAMP (Collaborative Harmonisation of the Profiling Methods of Amphetamine Type Stimulants) Work Package 4: Development of an International Common Database and Application of Statistical Treatments", Helsinki 2006
- (I22) 18th International Symposium on the Forensic Sciences The Australian and New Zealand Forensic Science Society (ANZFSS), **L. Dujourdy**, présentation orale : "Appraisal of various statistical tools for use in drug profiling", Perth 2005
- (I23) Séminaire/atelier : "Efficient Amphetamine Profiling by use of a Harmonised Method and a Common Database", Commission Européenne, Bruxelles 2005
- (I24) 17th Meeting of the International Association of Forensic Sciences (IAFS), **L. Dujourdy**, F. Besacier, présentation orale : "A quick and automated profiling method of cocaine HCl samples for tactical intelligence purposes", Hong-kong 2005
- (I25) 11th Meeting European Network Forensic Sciences Institute (ENFSI) Working group Drugs, **V. Dufey, L. Dujourdy**, F. Besacier, H. Chaudron, 2 posters : "A quick and automated profiling method of heroin samples for tactical intelligence purposes", "Cocaine profiling for strategic intelligence purposes, a cross-border project between France and Switzerland", Lisbonne 2005
- (I26) 14th International Forensic Science Symposium, poster : "Cocaine profiling for strategic intelligence purposes. A cross-border project between France and Switzerland", INTERPOL, Lyon 2004
- (I27) 3rd European Academy of Forensic Science (EAFS) triennial Meeting, **L. Dujourdy**, présentation orale : "Evaluation of links in heroin seizures" ; poster : "Chemometric analysis of RAMAN spectroscopy data on MDMA in "ecstasy" tablets", Istanbul 2003
- (I28) 16th International Symposium on the Forensic Sciences, P. Esseiva, F. Anglada, **L. Dujourdy**, F. Taroni, P. Margot, E. Du Pasquier, M. Dawson, C. Roux, P. Doble, poster : "Application of neural network in drug intelligence based on chemical profiling", Canberra 2002
- (I29) 16th meeting of the International Association of Forensic Sciences, 2 posters : **L. Dujourdy**, "Evaluation of links in heroin seizures" ; P. Esseiva, **L. Dujourdy**, F. Anglada, F. Taroni, P. Margot, "A methodology for illicit heroin seizures comparison in a drug intelligence perspective using large databases", Montpellier 2002
- (I30) 7th Meeting European Network Forensic Sciences Institute Working group Drugs, **L. Dujourdy**, 2 présentations orales : "Cosine function : a way to link heroin seizures" ; "Development of a harmonised pan-European method for the profiling of amphetamines : a presentation", Madrid 2001
- (I31) 2nd European Network of Forensic Science Institutes meeting, **L. Dujourdy**, présentation orale : "Data treatment in drug intelligence using heroin impurities and amphetamine impurities", Cracovie 2000
- (I32) Euro-Fillers 95, **L. Dujourdy**, J. P. Cohen Addad, présentation orale : "Kinetics of polydimethylsiloxane adsorption on silica aggregates. Effect of chain molecular weight and silica concentration", Mulhouse 1995

• Nationales

- (N1) Les 3^e Rencontres Francophones sur les Légumineuses (RFL3), J. Melendrez-Ruiz, I. Goisbault, J-C. Charrier, K. Pagnat, **L. Dujourdy**, G. Arvisenet, S. Chambaron, poster : "Une étude exploratoire en supermarché virtuel pour encourager les consommateurs à choisir des produits alimentaires à base de légumes secs : quel impact de scénarios et de nudges ? ", Angers, 2021
- (N2) XX^e Congrès Chimiométrie, SupAgro, C. Peltier, A. Câmara Jr., C. Sandt, F. Husson, **L. Dujourdy**, poster: "Geometrical patterns in PCA with spectral data. Why ?", Montpellier, 2019
- (N3) AGROSTAT 2018, C. Peltier, **L. Dujourdy**, présentation orale : "What is the better test to detect multivariate differences in large dimensional data?", Marseille, 2018
- (N4) XIX^e Congrès Chimiométrie, CNAM, C. Peltier, C. Modugno, A. Jossier, **L. Dujourdy**, H. Simonin, JM. Perrier-Cornet, poster : "IR-spectral analysis and chemometrics to study the response of Bacillus subtilis spores to high pressures", Paris, 2018
- (N5) Colloque Chimie et Expertise : sécurité des biens et des personnes, **L. Dujourdy**, présentation orale : "Outils et techniques de profilage des drogues", Maison de la Chimie – Paris 2014
- (N6) XVII^e congrès annuel de la Société Française de Toxicologie Analytique (SFTA), M.-H. Ghysel-Laporte, C. Robinot, F. Bouchet, M. Mancebo, S. Layec, F. Sengel-Lebreton, A. Daniel, **L. Dujourdy** , F. Besacier, poster : "Les produits stupéfiants consommés en 2008", La Rochelle 2009
- (N7) Journées TECHNOPOLICE, **L. Dujourdy**, présentation orale : "Le profilage chimique de stupéfiants", Gif/Yvette, 2009

Chapitre 1 : Curriculum vitae

- (N8) XV^e congrès annuel de la Société Française de Toxicologie Analytique (SFTA), F. Bévalot, **L. Dujourdy**, L. Magné, F. Besacier, C. Le Meur, Y. Gaillard, L. Fanton, D. Malicier, présentation orale: "Interprétation des concentrations en méprobamate dans l'humeur vitrée, la bile et la moelle osseuse: approche statistique", Paris 2007
- (N9) Forum Labo & Forum Biotech, **L. Dujourdy**, présentation orale: "Le cannabis "haut dosage", un produit à dormir debout", CNIT Paris - La Défense 2006
- (N10) XIII^e congrès annuel de la Société Française de Toxicologie Analytique (SFTA), S. Locicero, P. Esseiva, P. Hayoz, **L. Dujourdy**, F. Besacier, P. Margot, poster: "Cocaine profiling for strategic intelligence purposes. A cross-border project between France and Switzerland", Pau 2005
- (N11) Journée Club Échantillonnage et Analyse Rhône-Alpes (CERA), Applications statistiques - Étude de cas concrets, **L. Dujourdy**, présentation orale: "Statistiques lors de la recherche de stupéfiants", Lyon 2005.
- (N12) Physique en clips, **L. Dujourdy**, présentation orale: "Mélanges de polydiméthysiloxane et particules de silice. Adsorption compétitive et cinétique", Grenoble 1994
- (N13) Carrefours de la Fondation : transfert Recherche/Industrie, **L. Dujourdy**, JP. Cohen-Addad, poster: "Évolution des propriétés viscoélastiques des matériaux silicone/silice en fonction de la durée de stockage", Charbonnières 1994
- (N14) Sciences en fête, **L. Dujourdy**, présentation orale: "Démonstration des propriétés insolites des polymères", Grenoble 1993

Chapitre 2

Expériences en matière de recherche

2.1. Introduction

Ce mémoire me permet de faire un bilan d'une vingtaine d'années de recherche appliquée, d'abord dans le contexte particulier d'un laboratoire de police scientifique puis dans une école d'ingénieur.

Au-delà de la mise en œuvre de méthodes analytiques et statistiques, il sera proposé toute une méthodologie pour exploiter au mieux et fiabiliser des données d'analyses.

Cette thématique a commencé dans un environnement académique à l'université de Lausanne dans un institut de police scientifique durant 3 ans. Elle s'est ensuite développée pendant 13 ans dans un univers non universitaire où les enjeux et priorités du chercheur doivent s'adapter au contexte particulier d'aide à l'enquête judiciaire pour lequel une réponse est attendue parfois dans des délais très courts ; cette réponse pouvant être défendue devant un tribunal.

Exercer au sein d'un établissement public administratif du Ministère de l'Intérieur a donc façonné mes orientations avec pour leitmotiv l'interprétation des résultats comme preuve scientifique dans un cadre pénal.

Attirée par la transmission des connaissances et désireuse d'élargir mon champ de compétences, j'ai eu l'opportunité, grâce à un détachement, d'évoluer de nouveau dans un univers académique à Agrosup Dijon nouvellement renommé « l'Institut Agro Dijon » depuis son intégration en janvier 2022 dans l'Institut Agro (institut national d'enseignement supérieur pour l'agriculture, l'alimentation et l'environnement). Depuis 6 ans, j'explore les différentes facettes de l'analyse de données avec pour objectif de poursuivre ma compréhension du profilage physico-chimique.

2.2. Quelques définitions et éléments de contexte

2.2.1. Que signifie être chercheur dans un cadre pénal ?

On pourrait dire que c'est faire de la « **science forensique** », c'est-à-dire, appliquer une démarche scientifique et des méthodes techniques dans l'étude des traces qui prennent leur origine dans une activité criminelle, ou litigieuse en matière civile, réglementaire ou administrative ¹. Science forensique provient de la traduction de l'anglais « forensic science », rendu nécessaire par la confusion des termes et de leurs traductions qui désignent la contribution des sciences, en particulier des sciences de la nature, à la justice ². En français, on rencontre le terme de **criminalistique** qui est l'ensemble des techniques mises en œuvre par les services de la police et de la justice pénale pour établir la preuve d'un crime et pour identifier son auteur. On appelle généralement ce secteur du nom populaire de « police technique et scientifique (PTS) ». C'est également une discipline scientifique autonome s'articulant autour de la trace, vestige et témoin silencieux d'une action criminelle ou accidentelle. **La recherche scientifique en criminalistique a un rôle primordial pour toujours replacer l'étude de la trace à l'intersection de la science, de la police et du droit.**

Au quotidien, qu'est-ce que cela signifie ? C'est apporter des éléments de preuves c.-à-d. indices de nature scientifique, robustes, fiables et vérifiables pour participer à la réduction de l'incertitude des enquêteurs et des magistrats, tout au long de l'enquête. C'est donc mettre en place des méthodologies opérationnelles aux démarches d'innovation et de développement^a.

Les activités analytiques proposées par les laboratoires de police scientifique couvrent la majeure partie des champs de la criminalistique : identification de la personne (établissement de profils génétiques, portrait-robot génétique et recherche de traces papillaires), physique-chimie (microtraces de peintures, verres, fibres, résidus de tirs ...), toxicologie (composition et dosage de toxiques prélevés sur un corps lors d'une autopsie, ou alors de prélèvements en lien avec l'alcoolémie, les stupéfiants ou un empoisonnement), stupéfiants (identification et dosage des produits, recherche des produits de coupage, profilage physico-chimique), incendies et explosions (recherche des causes des incendies, d'hydrocarbures ou d'explosifs), balistique (examen des armes, munitions et trajectoires de tir).

Cela implique de travailler dans des délais fixés qui peuvent être très courts notamment dans le cas d'une demande de résultats lors d'une garde à vue.

Un ingénieur de PTS est recruté dans une discipline en particulier pour exercer des missions en premier lieu dans un laboratoire. Les ingénieurs participent aux analyses puis sont principalement chargés de l'exploitation des résultats analytiques et de la rédaction de rapports d'interprétation destinés aux services d'enquête. Dans tous les cas, leur rôle ne se limite pas à la seule partie scientifique. Ils doivent gérer les effectifs, contrôler les résultats, ils peuvent avoir à gérer le fonctionnement d'une section ou encore à exercer d'autres missions (formation, communication, R&D, etc.).

Ces éléments et ces contraintes exercent une influence sur la façon d'appréhender les activités de recherche et de développement et j'y reviendrai dans la suite de ce manuscrit.

Avant d'intégrer la police scientifique française, j'ai réalisé d'abord un stage postdoctoral à l'Institut de police scientifique et de criminologie (IPSC) de l'université de Lausanne sous la direction du professeur Pierre Margot, alors directeur de l'Institut. J'ai ensuite obtenu un poste contractuel de maître-assistante dans le domaine de l'analyse des stupéfiants. Ayant une formation en sciences physiques et en modélisation de comportements physico-physiques de matériaux polymères, je me suis concentrée sur l'analyse de données issues de techniques de séparation (chromatographie gazeuse et électrophorèse capillaire) et plus particulièrement à l'établissement d'empreintes chimiques, appelé profilage chimique, permettant de cartographier les filières d'approvisionnement.

Le Professeur Pierre Margot a été intronisé au Panthéon francophone de la criminalistique et il est mentionné que par sa vision de la recherche et de la science, il a fait rayonner l'École des sciences criminelles (anciennement IPSC) dans le monde entier. Il en a fait un lieu « central d'une réflexion sur la criminalistique, en tant que discipline scientifique autonome s'articulant autour de la trace, vestige et témoin silencieux d'une action criminelle ou accidentelle. Il a su mettre la recherche et les considérations scientifiques au cœur de la criminalistique, toujours dans une perspective replaçant l'étude de la trace à l'intersection du droit, de la police et de

^a Service National de Police Scientifique, « La Police scientifique : une activité essentielle pour la police nationale » dans <https://www.police-nationale.interieur.gouv.fr/Organisation/SNPS/Presentation>, consulté le 20/04/2022

la science »^b. Ce fut une chance et un honneur de le côtoyer et de travailler à ses côtés et c'est ce qui m'a permis de réussir le concours pour devenir ingénieur de PTS au laboratoire de police scientifique de Lyon du ministère de l'Intérieur. J'ai été affectée dans la section stupéfiants du laboratoire où j'ai continué à m'investir sur ce sujet bien que n'étant plus dans un environnement académique.

Structure de la police scientifique en Suisse et en France

La Suisse est structurée en 26 cantons qui sont les États fédérés de la Confédération suisse. Chaque canton a son propre mode de fonctionnement notamment en ce qui concerne la police scientifique.

Il existe 2 lieux de formation aux techniques de la police scientifique. Tout d'abord, l'École des Sciences Criminelles (ESC), rattachée à la Faculté de droit, des sciences criminelles et d'administration publique de l'Université de Lausanne, elle a été créée par le Professeur Rodolphe Archibald Reiss en 1909. Elle fut la première école de police scientifique au monde et demeure l'une des seules institutions en Europe à offrir une formation complète en science forensique. Elle dispose également d'un service d'expertises qui est en lien étroit avec la recherche, l'enseignement et la formation continue.

L'autre lieu est un service scientifique à Zurich qui existe depuis 1952.

En France, les fonctionnaires de la police scientifique sont tous des fonctionnaires d'État dépendant du ministère de l'Intérieur. Ils appartiennent à deux grandes directions : la Direction Générale de la Police Nationale (DGPN) et à la Direction Générale de la Gendarmerie Nationale (DGGN).

L'organisation de la police scientifique au sein de la Police nationale a évolué au cours des 20 dernières années. En janvier 2021, un grand changement est intervenu avec la création du Service national de police scientifique (SNPS), structure issue de la fusion de 2 entités de criminalistique au sein de la Police nationale, le SCPTS (Service central de la police technique et scientifique) et l'INPS (Institut national de police scientifique) qui est le réseau de laboratoires de police scientifique en France. Le siège du SNPS se situe à Lyon.

Les laboratoires de police scientifique (LPS) sont au nombre de cinq et se situent à Lille, Lyon, Paris, Toulouse et Marseille.

La gendarmerie dispose d'un laboratoire situé en région parisienne, il s'agit de l'Institut de Recherche Criminelle de la Gendarmerie Nationale (IRCGN).

2.2.2. Qu'entend-on par profilage physico-chimique ?

Suivant le domaine, on peut trouver plusieurs définitions du profilage physico-chimique. Par exemple, dans le domaine pharmaceutique, on parlera de profilage d'impuretés, celui-ci vise à identifier et à quantifier des composants spécifiques présents à de faibles niveaux, généralement inférieurs à 1% et idéalement inférieurs à 0,1%. Les impuretés étant des résidus indésirables qui se forment pendant ou après le déroulement de la réaction de synthèse. Les résidus peuvent être des produits chimiques inorganiques, organiques ou des solvants résiduels qui définissent la qualité et les propriétés d'un produit spécifique, par exemple l'efficacité d'un ingrédient pharmaceutique actif (API) 3.

^b François Beaudoin B.A., historien de l'Association québécoise de criminalistique, « Pierre Margot : les sciences forensiques, l'interdisciplinarité requise », <https://criminalistique.org/assets/files/PantheonPMargot.pdf>, consulté le 20/04/2022

On peut rencontrer également l'appellation d'empreinte ou de signature chimique et dans ce cas, il s'agit d'un modèle (pattern) unique indiquant la présence d'une ou plusieurs molécules particulières (cibles), obtenue à l'aide de techniques analytiques spécialisées comme la spectroscopie et/ou la spectrométrie de masse et utilisée pour caractériser un polluant, un médicament, un contaminant ou un autre produit chimique dans un échantillon testé.

L'obtention d'une empreinte chimique est réalisée en intégrant la chimie, les sciences analytiques et les mathématiques, on trouve également le terme de chimiométrie.

La génération de profils ou d'empreintes d'impuretés dans des formats qui permettent d'utiliser les données à différentes fins, notamment pour la détermination de l'origine, de la variété, des conditions agroclimatiques, du lien entre les lots ... constitue le processus de comparaison.

Le profilage peut être vu comme un cas particulier d'un procédé global d'analyses dans lequel l'analyse statistique est présente tout au long de la chaîne (figure 1) ; l'objectif étant d'évaluer les risques d'erreur à chaque étape.

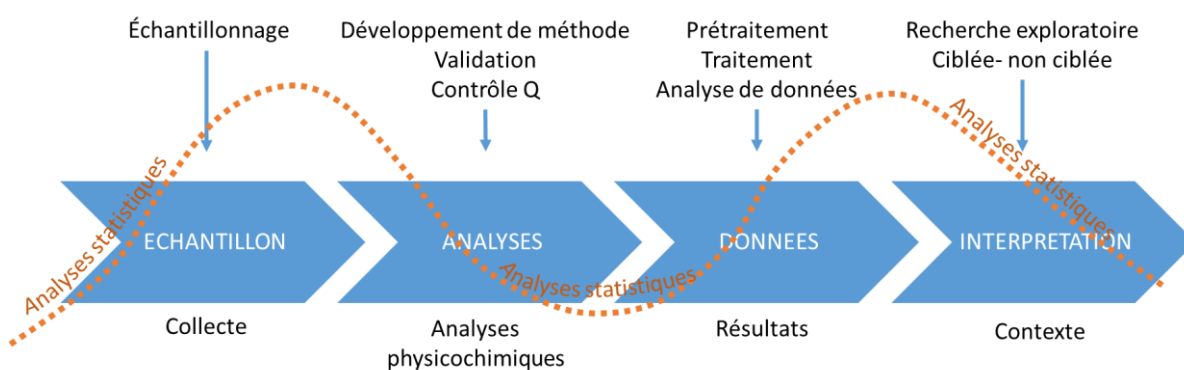


Figure 1 : Procédé global d'analyses.

Dans le domaine de l'analyse des drogues illicites, la combinaison du profilage des impuretés et de la comparaison de profils à des fins de renseignement (stratégique, par ex. détermination des routes d'approvisionnement, tactique, par ex. lien entre 2 échantillons saisis sur 2 personnes différentes) et de preuve est connue sous le nom de profilage des drogues illicites^c.

Le premier stade de la démarche est consacré au développement de méthodologies analytiques permettant l'extraction de la signature d'un produit illicite. Le deuxième stade est directement axé sur l'application des résultats au service des forces de police et de la justice.

2.2.3. Qu'est-ce la chimiométrie ?

Il existe plusieurs définitions, voici celle donnée dans l'ouvrage de M. Otto⁴ : « discipline chimique qui utilise des méthodes mathématiques et statistiques (a) pour concevoir ou sélectionner des procédures de mesure et des expériences optimales, et (b) pour fournir un maximum d'informations chimiques en analysant les données chimiques. ».

^c Définition du profilage des drogues, European Network of Forensic Science Institute, réunion annuelle du groupe de travail sur les drogues, Cracovie, Pologne 2006

On trouve également : « la chimiométrie est une branche de la chimie analytique qui utilise les outils mathématiques et statistiques dans le but d'obtenir le maximum d'informations à partir des mesures réalisées en laboratoire ou sur des procédés industriels »⁵.

Dans son ouvrage Brereton⁶ décrit les relations entre la chimiométrie et différentes disciplines (figure 2).

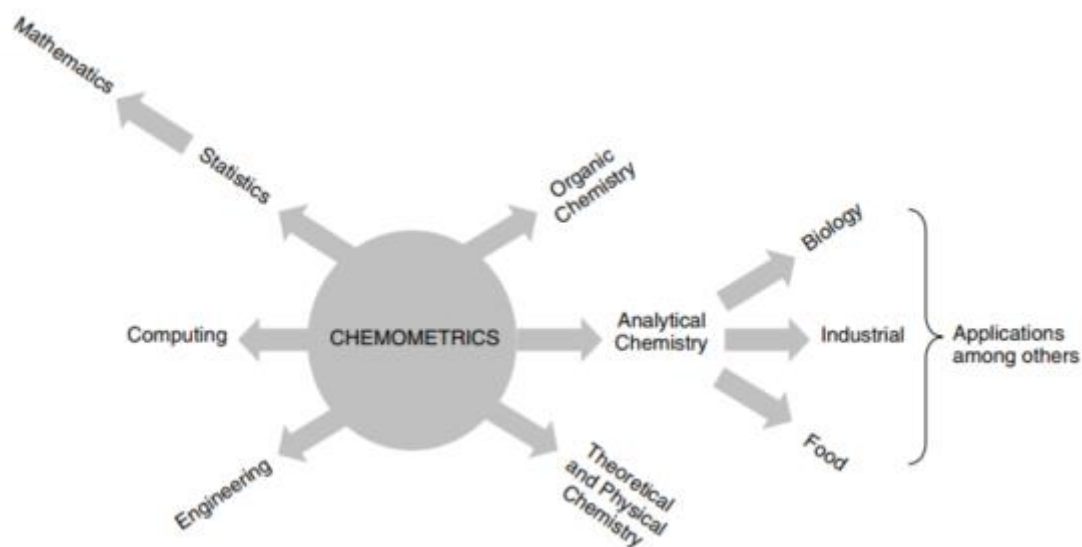


Figure 2 : Lien entre la chimiométrie et d'autres disciplines d'après Richard G. Brereton, "Chemometrics: Data Analysis for the Laboratory and Chemical Plant", John Wiley & Sons, Ltd, 497 pages, 2003

Toute expérimentation a au moins un objectif et une stratégie est construite pour l'atteindre. Parmi les étapes de cette construction, les mesures issues d'appareils d'analyse figurent en bonne place^d. Concernant les mesures (variables quantitatives), voici quelques questions qu'on peut se poser :

- Quelle confiance peut-on accorder à ces mesures ?
- Comment peut-on exprimer cette confiance d'une façon scientifiquement fiable ?
- Ces mesures sont-elles justes ? Vraies ?
- Aurait-on pu avoir de meilleurs résultats ? Comment ?
- Une autre stratégie donnerait-elle les mêmes résultats ?

Ces questions font partie intégrante de la chimiométrie.

La chimiométrie ne s'occupe pas « a priori » des incertitudes calculées à partir des caractéristiques des appareils de mesures qui est plutôt du domaine de la métrologie. Mais on verra par la suite qu'elles interviennent dans la validation d'une méthode d'analyse quantitative.

2.2.4. Qu'est-ce qu'une donnée ?

Il existe bien des définitions, mais celle qui, je pense, correspond assez bien au cadre dans lequel ce mémoire fait référence est celle de Pascal Rivière, chef de l'Inspection générale à l'Insee : « Toute donnée se caractérise par un vaste faisceau de conventions (sémantique, nomenclatures, formats, etc.), et par l'infrastructure de

^d H. Huser, « Introduction à la chimiométrie. Mesures et traitement statistique », Séminaire Echantillonnage et Mesure en environnement – Egletons 2011

connaissances dans laquelle elle s'inscrit, impliquant des choix qui n'ont rien de neutre. Une donnée se révèle aussi dépendante de l'environnement qui lui a donné naissance, et des processus productifs qui l'utilisent. On constate alors que les données ne sont pas pures et parfaites, ne vont pas de soi : paradoxalement, les données ne sont pas données. Il faut les construire, les prendre. Elles requièrent en amont un travail de modélisation, d'abstraction, de spécification des concepts puis des domaines, avant d'imaginer produire des valeurs. »⁷.

Interpréter des données pourrait se résumer à globalement acquérir des connaissances sur un phénomène c'est-à-dire à savoir ce que les données peuvent bien nous apprendre dans un cadre spécifique. Toutefois, chaque méthode statistique va révéler un aspect des données sans forcément fournir un point de vue global, voire une vérité sur le phénomène étudié. **Il est donc primordial de maîtriser la chaîne de production de celles-ci et c'est ce que je me suis employée à faire durant ces années.**

Le plan de la suite de ce mémoire reprendra les étapes décrites dans la figure 1 et ne sera pas chronologique. Dans le paragraphe 2.3 seront abordés la caractérisation des stupéfiants et le profilage puis le paragraphe 2.4 abordera les questions d'échantillonnage. Le paragraphe 2.5 sera consacré à la production de résultats dans une démarche d'assurance qualité. Le traitement de ces résultats et leur interprétation dans un cadre donné seront décrits dans le paragraphe 2.6. Enfin, le chapitre 3 sera consacré à un changement de cap et aux nouvelles perspectives qu'il offre en lien avec mon thème de recherche.

2.3 : Caractérisation des stupéfiants et profilage

2.3.1 Éléments de contexte

a) Qu'est-ce qu'une drogue ?

Définition donnée par la Mission Interministérielle de Lutte contre les Drogues et les Conduites Addictives (MILDECA) publiée le 07/08/2015 sur son site web^e : on appelle « drogue » toute substance psychotrope ou psychoactive qui perturbe le fonctionnement du système nerveux central (sensations, perceptions, humeurs, sentiments, motricité) ou qui modifie les états de conscience.

Une drogue est un produit susceptible d'entraîner une dépendance physique et/ou psychique. Les dangers ou risques d'une substance dépendent de nombreux facteurs : l'âge et le sexe du consommateur, le mode de consommation (ingestion, « sniffing », injection, inhalation de vapeurs), la fréquence à laquelle il consomme, etc. En général, parmi les drogues, on peut distinguer celles qui semblent apaiser, telles que l'alcool, les tranquillisants (médicaments), le cannabis ou les opiacés ; celles qui stimulent, telles que la cocaïne, l'ecstasy, les amphétamines ou les cathinones de synthèse (ex. famille des méthylméthcathinones MMC) ; celles qui provoquent des hallucinations, telles que le LSD ou les champignons hallucinogènes. Ces effets peuvent varier selon la dose consommée.

Le terme de drogues recouvre aussi bien les **drogues illicites** comme le cannabis, l'héroïne ou la cocaïne que les **drogues licites** comme la nicotine, l'alcool ou les médicaments.

^e Site web : <https://www.drogues.gouv.fr/comprendre/l-essentiel-sur-les-addictions/qu-est-ce-qu-une-drogue> consulté le 21/04/2022

Le terme « **stupéfiant** » a une **signification juridique** : c'est le terme utilisé pour désigner les **substances psychoactives interdites**, inscrites sur une liste des « stupéfiants »^f. Certains médicaments font l'objet d'une réglementation particulière, comme la morphine. Le terme « **psychotropes** » a également une signification juridique : c'est le terme utilisé pour désigner les substances psychoactives réglementées, car appartenant à la catégorie des substances vénéneuses.

L'objectif du classement (cf. ci-après) des produits est de limiter l'usage des stupéfiants et des psychotropes aux seules fins médicales et scientifiques et d'encadrer leur utilisation afin d'éviter tout abus ou détournement vers l'usage illégal ou le trafic, qui constituent des infractions pénales passibles de peines de prison et d'amendes. La liste des « stupéfiants » est régulièrement mise à jour en fonction de la connaissance des substances dangereuses disponibles sur le marché.

Depuis une vingtaine d'années, de nouvelles substances psychoactives naturelles ou synthétiques (NSP) apparaissent sur le marché. Fabriquées en laboratoire, elles imitent les effets du cannabis, de la cocaïne, de la MDMA / ecstasy, des amphétamines ou d'autres drogues illégales. Les substances les plus courantes sont les cathinones de synthèse comme la 3-MMC et les cannabinoïdes de synthèse^g comme le MDMB-4en-PINACA^h.

Elles peuvent provoquer des intoxications parfois mortelles. Dès que la preuve de leur caractère dangereux est apportée, leur classement dans la liste des stupéfiants ou des psychotropes est demandé, soit au niveau national, soit au niveau international. Un classement générique (par famille, arrêté du 27 juillet 2012) a été introduit pour les cathinones, les cannabinoïdes, les NBOMe (abréviation de N-méthoxybenzyle) et les benzofuranes.

b) Classification^h des drogues par leurs effets

On distingue les substances psychoactives **d'origine naturelle** : extraite d'une plante ou d'un champignon, à l'état quasi brut ou retraité chimiquement des substances **d'origine synthétique** : totalement fabriquée en laboratoire à partir de produits chimiques.

- « **dépresseur** du système nerveux central » : elles agissent sur le cerveau en ralentissant certaines fonctions ou sensations. Un ralentissement de la fonction respiratoire et l'endormissement sont souvent des effets secondaires de ces drogues. Exemples : famille des opiacés (dérivés de l'opium, plante cultivée en Asie) : héroïne, morphine, codéine et les barbituriques, hypnotiques, neuroleptiques.

- « **stimulant** du système nerveux central » : elles accroissent les sensations et certaines fonctions organiques comme le rythme cardiaque ou encore la sensation d'éveil. Leur action « stimulante » est souvent suivie d'un contre-coup avec, par exemple, des sensations inverses de fatigue et d'irritabilité. Exemples : amphétamines, cocaïne, caféine, produits dopants.

^f Site web : <https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000533085/> consulté le 21/04/2022

^g Roue interactive permettant de rendre la chimie des cannabinoïdes synthétiques plus facilement compréhensible <https://www.emcdda.europa.eu/topics/pods/synthetic-cannabinoids>, site web consulté le 24/09/2022

^h Classification établie par Y. Pélacier et J. Thuillier en 1991

« **hallucinogène/perturbateur** » : elles modifient les perceptions visuelles, auditives et corporelles. Ces modifications sont très dépendantes du contexte et de la personne qui utilise de telles drogues. Exemples : LSD, kétamine, psilocybine, cannabis.

c) Contexte juridique dans lequel sont analysées les saisines

Une « saisine » est le fait de recourir à une juridiction ou une entité de médiation afin de lui soumettre un litige en vue de la résolution de celui-ci. Une saisine est donc un acte volontaire et écrit du requérant.

Classement des stupéfiants et des psychotropes au niveau international

L'Organisation des Nations Unies a établi plusieurs conventions internationales afin de contrôler les stupéfiants et les psychotropes au niveau mondial.

- La Convention unique de 1961 sur les stupéfiants, amendée par le protocole de 1972, est de compétence nationale.
- La Convention de 1971 sur les substances psychotropes, est de compétence nationale.
- La Convention de 1988 contre le trafic illicite de stupéfiants est de la compétence communautaire.

L'OMS est responsable de l'évaluation des substances relevant des Conventions de 1961 et 1971.

L'OICS (Organe international de contrôle des stupéfiants) est chargé d'évaluer les substances relevant de la Convention de 1988.

Deux principes découlant de ces Conventions sont à retenir. D'une part, l'OMS se doit d'être en permanente collaboration avec les institutions scientifiques, les services de santé, les organismes de réglementation, les autorités sanitaires, de police, les organismes intergouvernementaux et les organisations non gouvernementales (ONG). D'autre part, l'OMS s'appuie sur les compétences de son Comité d'experts de la pharmacodépendance.

Classement des **stupéfiants** selon les conventions internationales

La Convention de 1961 sur les stupéfiants classe ces substances en 4 Tableaux :

- Le Tableau I concerne les abus et les effets nocifs comparables à la morphine, la cocaïne ou le cannabis
- Le Tableau II concerne les risques comparables à la codéine ou au dextropropoxyphène
- Le Tableau III rassemble les préparations des substances classées dans les Tableaux I et II qui sont sans risque d'abus ni d'effets nocifs ainsi que les substances non aisément "récupérables" ou extractibles
- Le Tableau IV fait état des substances du Tableau I ayant un potentiel d'abus fort et des effets nocifs importants, sans valeur thérapeutique notable

Classement des **psychotropes** selon les conventions internationales

L'inscription d'une substance à un Tableau de la Convention de 1971 repose sur l'évaluation du potentiel de dépendance et de l'activité sur le système nerveux central ou de l'existence d'un potentiel d'abus et d'effets nocifs comparables à ceux des substances déjà inscrites ainsi que sur l'évaluation des risques de santé publique ou sociaux encourus.

Comme pour les stupéfiants, les psychotropes sont classés au niveau international en 4 Tableaux :

- Le Tableau I : substances dont le potentiel d'abus présente un risque grave pour la santé publique et dont la valeur thérapeutique est faible
- Le Tableau II : substances dont le potentiel d'abus présente un risque sérieux pour la santé publique et ayant une valeur thérapeutique considérée comme faible à moyenne
- Le Tableau III : substances ayant un potentiel d'abus présentant un risque sérieux pour la santé publique, mais possédant une valeur thérapeutique moyenne à grande
- Le Tableau IV : substances avec un potentiel d'abus présentant un risque faible pour la santé publique, mais présentant une valeur thérapeutique faible à grande

Transposition en droit français

Le processus national de classement d'une substance donnée se base sur les éléments d'appréciation élaborés par l'OMS ainsi que sur les données fournies par le système national d'évaluation de la pharmacodépendance.

Les Centres d'Évaluation et d'Information sur la Pharmacodépendance (CEIP) recueillent les données de pharmacodépendance, les évaluent ensemble lors des réunions en Comité Technique et proposent de les traiter ou non en Commission Nationale des Stupéfiants et Psychotropes.

Le Conseil National des Professions de Santé (CNSP), au vu des travaux d'expertise présentés, donne son avis au Directeur général de l'Agence Nationale de Sécurité du Médicament et des produits de santé (ANSM) (par l'intermédiaire du Département Stupéfiants et Psychotropes) qui en tiendra compte pour prendre des mesures directement applicables ou le transmettra au ministre chargé de la santé pour suite à donner.

Classement français des stupéfiants

Par l'arrêté du 22 février 1990 modifié, le classement des stupéfiants au niveau international a été transposé ainsi en droit français.

La liste des substances classées comme stupéfiants au niveau national comprend 4 annexes :

- Les annexes I et II correspondent aux tableaux I et IV de la Convention Internationale sur les stupéfiants de 1961.
- L'annexe III comprend les substances des Tableaux III et IV et certaines substances des tableaux I et II de la Convention Internationale sur les psychotropes de 1971.
- L'annexe IV est constituée de substances psychoactives non classées au plan international et de certains précurseurs.

Classement français des psychotropes

La transposition a été effectuée par un arrêté en date du 22 février 1990 modifié. Le classement français des psychotropes se fait selon 3 parties :

- La première partie correspond aux tableaux III et IV de la Convention internationale sur les psychotropes.
- La deuxième partie est composée des préparations de substances classées comme stupéfiants en France.
- La troisième partie est un classement à titre français c'est-à-dire qu'elle regroupe des substances non classées au niveau international.

Les mesures de contrôle

- Concernant la fabrication et le commerce : les établissements pharmaceutiques fabriquant des produits contenant des substances classées comme stupéfiants ou psychotropes doivent y être autorisés par l'ANSM. Ces

établissements ont l'obligation de tenir un ordonnancier ainsi qu'un registre spécial des entrées et sorties de ces produits. Ces obligations s'inscrivent dans un souci de traçabilité.

- Concernant le commerce extérieur, toute importation ou exportation de ces substances doit faire l'objet d'un permis délivré par l'ANSM.
- S'agissant de la distribution, les grossistes répartiteurs sont également soumis à un régime d'autorisations délivrées par l'ANSM.
- De plus, les fabricants et les distributeurs sont tenus de tenir un état récapitulatif annuel indiquant les quantités reçues, les quantités utilisées, les quantités cédées et le stock en fin d'année.
- La liste des substances qui sont placées sous contrôle est établie par l'arrêté du 22 février 1990 modifié.
- Évaluation des besoins et statistiques soumis à approbation à l'OICS.

Les dispositions pénales

Les peines applicables :

- Concernant les stupéfiants : articles L. 3421-1 à L.3421-4 du Code de la santé publique.
- Concernant les substances vénéneuses : article L. 5432-1 du Code de la santé publique.
- Concernant les substances dangereuses : articles L. 1343-2 et L. 1342-4 du Code de la santé publique.

L'injonction et la prise en charge sanitaire des personnes toxicomanes :

- L.3412-1 et suivants du Code de la santé publique.

C'est donc dans un **contexte** non seulement **légal** avec des acteurs de la chaîne judiciaire (police, gendarmerie, magistrature), mais également **sanitaire** (Agences de santé publique) que les **analyses** de substances supposées illicites **sont réalisées**.

2.3.2. Objectifs des analyses

Que ce soit en Suisse ou en France, les objectifs étaient et restent les mêmes à savoir l'analyse de milliers de prélèvements chaque année.

Les grandes familles de produits analysés sont en volume d'affaires (l'ordre peut varier certaines années) : la cocaïne, le cannabisⁱ, l'héroïne, les amphétamines et dérivés. Les nouveaux produits de synthèse (NPS) sont apparus dans les années 2000. Les premières identifications en France de molécules considérées comme étant des NPS ont été réalisées entre 2007 et 2008 et n'ont cessé de se diversifier en termes de familles chimiques et d'augmenter en nombre d'identifications^j.

Les autres stupéfiants tels que le LSD (diéthylamide de l'acide lysergique) et le GHB (acide 4-hydroxybutanoïque ou gamma-hydroxybutyrate),... sont moins fréquemment rencontrés. Pour le GHB, c'est plutôt son précurseur synthétique, le GBL, (gamma-hydroxybutyrolactone) qui est rencontré. La kétamine n'est actuellement plus retrouvée.

ⁱ Le nouvel arrêté réglementant le CBD en France, publié au JO du 31 décembre 2021 interdit la vente et la détention de fleur, mais autorise désormais (décision du Conseil de l'État du 24 janvier 2022) un taux de 0,3% de THC (Δ^9 -hydrocannabinol)

^j Observatoire Français des Drogues et Toxicomanies (OFDT), site web : <https://www.ofdt.fr/produits-et-addictions/de-z/nouveaux-produits-de-synthese/>, consulté le 21/04/2022

Les produits de coupage ainsi que les médicaments dont l'usage est détourné du marché légal sont également analysés.

Le personnel d'une section stupéfiants peut être amené à se déplacer pour procéder à des échantillonnages sur les grosses saisies, des prélèvements de traces dans des véhicules, appartements, sur des billets de banque..., des assistances lors d'investigation de laboratoires clandestins.

L'analyse d'une substance supposée illicite s'opère en 3 étapes :

- 1) une étape qualitative d'**identification**
- 2) une étape quantitative de **dosage**
- 3) une étape de **détermination de rapprochements** entre saisies

Il existe des recommandations au niveau international pour les 2 premières étapes. Pour la dernière partie, des directives sont proposées.

Les organismes proposant des documents sont les suivants :

- au niveau européen : le réseau européen des instituts de science forensique (European Network of Forensic Science Institutes - ENFSI) et notamment le groupe de travail sur les substances illicites (Drugs Working Group - DWG),
- au niveau international : le groupe de travail scientifique pour l'analyse des drogues saisies (Scientific Working Group for the Analysis of Seized Drugs - SWGDRUG)
- au niveau international non gouvernemental : L'Office des Nations Unies contre la drogue et le crime (United Nations Office on Drugs and Crime)

a) Étape 1 : l'identification

Le **résultat de cette 1^{re} étape aura pour effet la mise en cause ou non d'une personne** pour infraction liée à la législation sur les stupéfiants (ILS), que ce soit pour usage (Code de la santé publique, articles L11110-1 à L6441-1), ou pour trafic (Code pénal, articles 222-34 à 222-43-1). Notamment, l'article 222-37 du Code pénal réprime le transport, la **détention**, l'offre, la cession, l'acquisition ou l'emploi illicite de stupéfiants avec une **peine de dix ans** et 7 500 000 euros d'amende. Ces dispositions concernent la France. En Suisse, c'est la loi fédérale sur les stupéfiants (LStup ; RS 812.121) et les substances psychotropes et ses ordonnances qui s'appliquent.

C'est **étape est donc cruciale**. Elle implique que l'identification fiable et scientifiquement fondée d'une drogue ou d'un produit chimique dépend de l'utilisation d'un schéma analytique, c'est-à-dire une combinaison de techniques sélectionnées et appropriées par des analystes compétents dans le cadre d'un processus dont la qualité est contrôlée. Elle exige une sélection globale de techniques sur lesquelles les méthodes et procédures validées sont appliquées en laboratoire par du personnel formé et habilité^k.

Les techniques incorporées dans le schéma analytique peuvent être classées en 3 catégories en fonction du niveau de sélectivité qu'elles atteignent (cf. figure 3). Un schéma analytique approprié doit atteindre un niveau de sélectivité suffisant pour permettre une conclusion scientifiquement étayée et adaptée à la juridiction et aux protocoles du laboratoire.

^k Recommendations from the Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG), version 8, © SWGDRUG 2019-June-13, <https://www.swgdrug.org/index.htm>, consulté le 22/04/2022

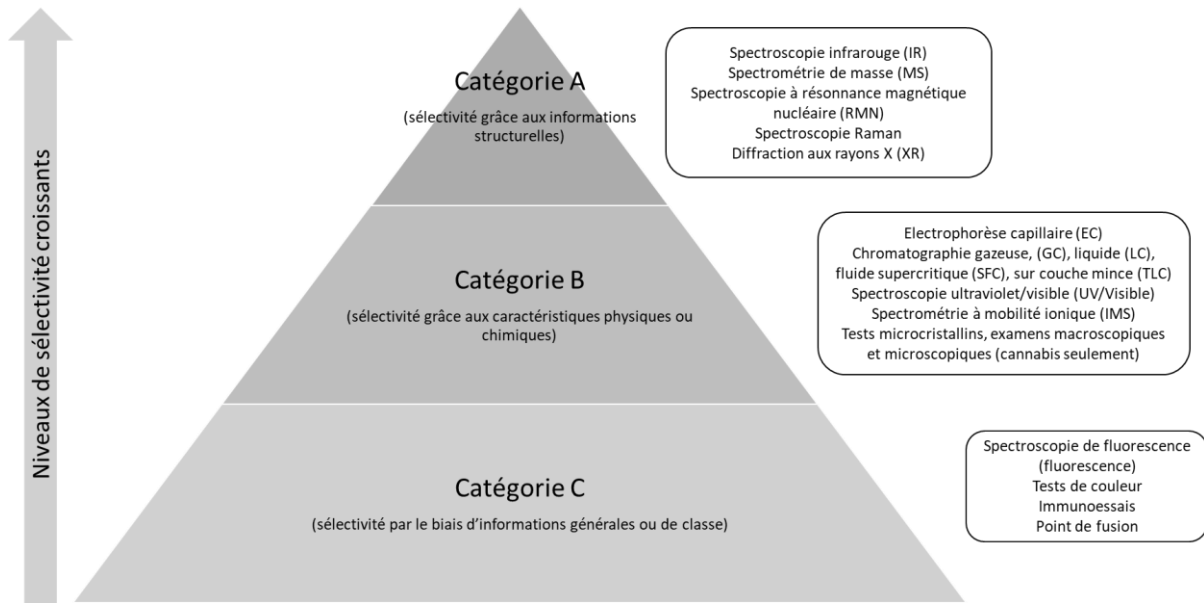


Figure 3 : Catégories des techniques analytiques rencontrées dans l'analyse de stupéfiants.

L'étape d'identification dans la plupart des laboratoires « modernes » peut être décrite par le schéma suivant (figure 4) :

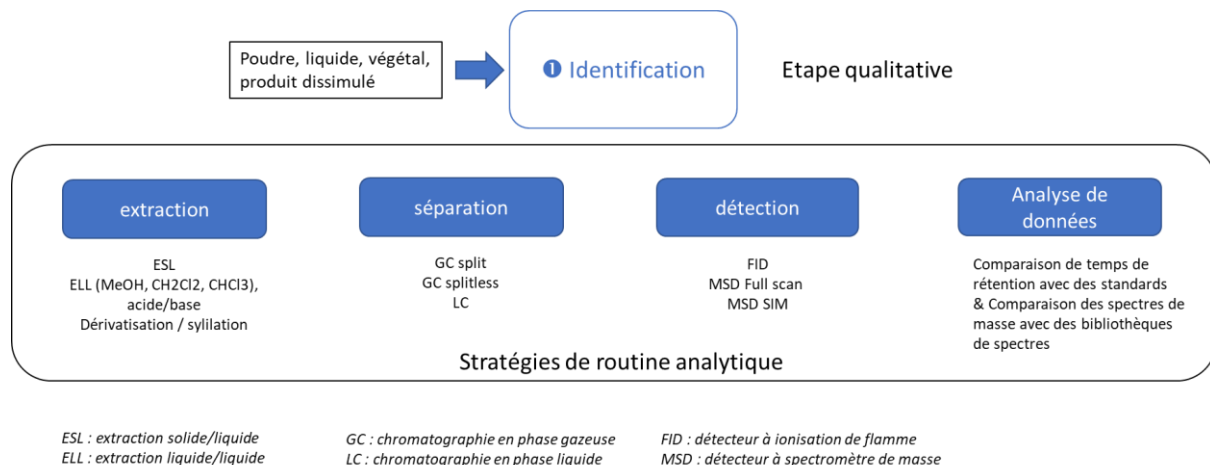


Figure 4 : Processus d'identification d'une substance supposée illicite

Des techniques de catégories C peuvent être utilisées en amont de cette étape, soit sur le terrain, notamment avec les tests à réactions colorimétriques ⁹, soit en laboratoire.

Suivant les cas, les stratégies visant à récupérer les composés d'intérêt peuvent être relativement longues et complexes, de même en cas de recherche de traces de stupéfiants sur des supports (couteaux, balances, billets de banque ...). Il est à noter toutefois que c'est cette étape qui est réalisée lorsqu'une demande est faite dans le cadre d'une garde à vue (qui est initialement de 24h, prolongeable jusqu'à 72h dans le cas de suspicion de trafic de stupéfiants). Il est alors nécessaire de recourir à de l'échantillonnage lorsque dans la saisie il y a de nombreux spécimens.

À l'issue de cette 1^{re} étape, le laboratoire détermine le classement des substances identifiées et commente la présence éventuelle de produits de coupage. Ces derniers peuvent être de 2 sortes, des diluants comme des sucres (sans effet psychoactif) et

des adjuvants comme des substances médicamenteuses (avec des effets psychoactifs, recherchés ou non).

L'apparition de toute nouvelle substance psychoactive fait l'objet d'un bulletin d'alerte rapide auprès de l'OFDT.

La forme chimique de certains stupéfiants est également déterminée. En effet, celle-ci conditionne son mode d'administration, par exemple, l'héroïne sous forme de sel peut être directement utilisée en injection intraveineuse alors que l'héroïne base sera plutôt consommée par inhalation des fumées (« chasser le dragon » ou « chasing the dragon ») ou prise ¹⁰. Néanmoins, l'héroïne base peut également être utilisée en injection intraveineuse après solvatation de la forme base dans l'eau grâce à un acide, généralement l'acide citrique (présent dans le jus de citron) et en la faisant chauffer. Elle renseigne également sur son marché : l'héroïne sous forme base libre est la forme majoritairement rencontrée en Suisse et en France. La cocaïne sous forme de sel chlorhydrate est la forme la plus consommée en France métropolitaine (dans les DOM hormis la Réunion, c'est le crack, sous forme base libre, qui est rencontré principalement).

En ce concerne l'héroïne, il est possible, par exemple, d'établir un lien entre la couleur et la forme chimique : les couleurs blanc et blanc cassé sont plutôt caractéristiques des héroïnes sous forme HCl, les couleurs allant du beige au marron pour l'héroïne sous forme de base libre ¹¹. Elles suivent par conséquent assez fidèlement l'évolution des formes chimiques illustrée à la figure 5.

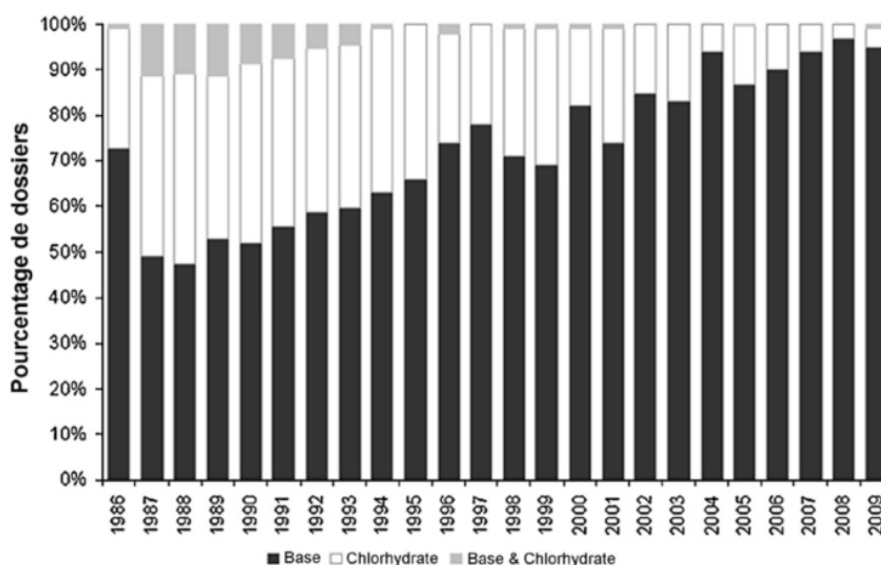


Figure 5 : Évolution de la forme chimique au cours du temps extrait de « L'héroïne saisie en France. Données statistiques issues de la base nationale des laboratoires de police scientifique » ¹¹

La cocaïne est principalement trouvée sous forme de poudre de couleur blanche, blanc cassé quelle que soit sa forme chimique. Des mélanges base et sel chlorhydrate correspondent à une mauvaise transformation de la cocaïne base en sel HCl ¹².

Les produits de coupage rencontrés dans la plupart des stupéfiants sont monitorés dès cette 1^{re} étape. L'ajout de ces substances obéit à différents critères comme la couleur, la granulosité, etc. Lorsque les substances sont inertes, on retrouve essentiellement des sucres. En revanche, lorsqu'elles sont pharmacologiquement actives, il s'agit de médicaments qui peuvent avoir différentes actions : servir à potentialiser le ou les effet(s) du produit stupéfiant ou au contraire à combattre ses

effets « négatifs » lors de la phase de « descente ». Enfin, la présence de certains médicaments reste parfois inexplicée. Leur identification présente un grand intérêt, car non seulement cela permet d'alerter ou d'éclairer les autorités sanitaires sur la toxicité des drogues en circulation (action pharmacologique spécifique au produit de coupage, risque d'overdose quand le produit est peu coupé, etc.), mais cela fournit aussi aux services répressifs des informations complémentaires pour remonter des filières : la nature et la concentration de ces produits étant souvent les mêmes pour une même source d'approvisionnement¹¹⁻¹⁴.

b) Étape 2 : le dosage (la quantification)

La **2^e étape** consiste à doser le ou les principes actifs et les principaux produits de coupage (figure 6).

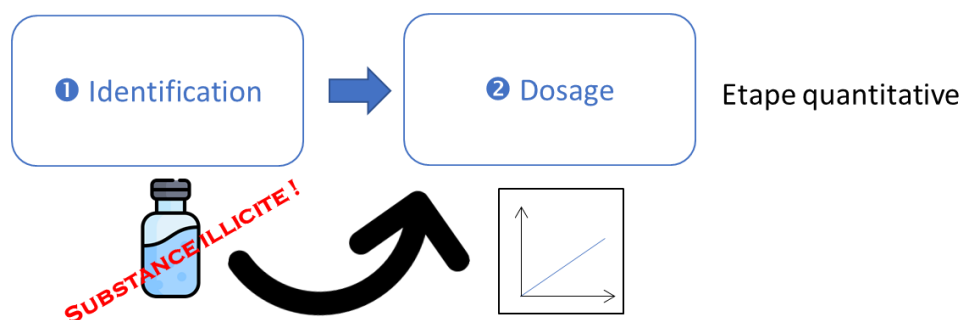


Figure 6 : Étape de dosage

En Suisse, la peine peut dépendre de la quantité. Par exemple, pour le cannabis, la consommation et la possession d'une quantité de cannabis inférieure à 10 grammes sont désormais considérées comme une simple infraction et passibles d'une amende d'ordre de 100 francs pour les consommateurs adultes (art. 19b, al. 2, LStup, en vigueur depuis le 1^{er} oct. 2013).

En France, le texte de loi ne précise pas la quantité à partir de laquelle on retient pour la détention de stupéfiants, il s'agit de considérer qu'une personne interpellée détenant de très petites quantités de stupéfiants peut se voir sanctionnée en vertu de l'article 222 – 37 du Code pénal.

Néanmoins, le dosage s'avère très utile pour « avoir une idée » du niveau de trafic auquel appartient la drogue. Par exemple, une héroïne base inférieure à 15% pourra être considérée comme une héroïne de rue (la teneur moyenne de l'héroïne saisie en France en 2021 est de 20.3%).

Afin de pouvoir fournir des résultats rapidement, les étapes d'identification et de quantification ont été fusionnées pour la plupart des analyses (en particulier en Suisse). Une seule méthode chromatographique (GC) sur colonne courte a été mise au point pour s'appliquer à tous les spécimens sous forme de poudre. La méthode consiste à peser précisément une certaine quantité de poudre, la solubiliser dans 1 ml d'une solution contenant un étalon interne (2 (LPS69 - France) ou 3 (IPS - Suisse) prises d'essai par spécimen sont réalisées et une moyenne est calculée). L'utilisation d'un étalon

¹ <https://www.ofdt.fr/produits-et-addictions/de-z/heroine-et-autres-opiaces/>

interne permet de s'affranchir de l'imprécision concernant le volume injecté, mais aussi de l'instabilité éventuelle de certains paramètres de l'appareil entre plusieurs injections successives.

Pour le cannabis, la méthode basée également sur une solution contenant un étalon interne permet de doser les 3 principaux cannabinoïdes : Δ^9 -tétrahydrocannabinol (Δ^9 -THC), cannabidiol (CBD) et cannabinal (CBN). Elle s'applique à tous les spécimens de cannabis (résine, herbe, huile).

Cette étape de quantification est importante, car elle permet de disposer pour tous les acteurs intervenant dans la prévention ou dans la répression du trafic des substances stupéfiants de disposer de données fiables sur la pureté des produits circulant sur le territoire national. De plus, l'évolution de celle-ci au cours du temps peut révéler certaines tendances en matière de consommation, mais aussi de disponibilité de tel ou tel produit. Dans certains cas, ces données permettent de renseigner le niveau de trafic auquel appartient la drogue (rue, grossiste ou laboratoire).

Toutes les informations issues des analyses sont enregistrées dans une base de données nationale STUPS® (Système de traitement uniformisé des produits stupéfiants), créée en 1986 à Lyon et maintenant partagée entre les 5 laboratoires de police scientifique (LPS) et l'institut de recherche criminelle de la gendarmerie (IRCGN). Elle permet de fournir une image représentative des produits stupéfiants saisis et analysés en France. Des statistiques sont éditées et fournies annuellement aux services sociaux et sanitaires (Observatoires français des drogues et des tendances addictives – OFDT - et Observatoire européen des drogues et toxicomanies - EMCCDA) et répressifs (Office anti-stupéfiants (OFAST), Mission de lutte antidrogue, etc.). Elle regroupe donc des informations issues des analyses, mais également des éléments de contexte comme les dates, lieu, provenance, destination des saisies (quand disponibles). Elle contient aujourd'hui des milliers de fiches de différents types de stupéfiants totalisant environ 10 millions d'entrées.

Ces données ont servi également à étudier la « puissance » du cannabis circulant en France sur une période de 25 ans ¹⁵. Cette thématique est importante d'un point de vue sanitaire et les résultats ont été repris dans 49 autres publications internationales.

c) Étape 3 : les comparaisons

Dans cette étape, on compare les saisies d'un stupéfiant en particulier entre elles.

Le profilage des drogues illicites est le procédé d'extraction de composés cibles à partir des saisies. Sa finalité est de fournir des informations objectives utiles pour décrypter l'organisation du trafic de stupéfiants (les réseaux de distribution) à des fins de renseignement pour les services d'enquête. Elles peuvent aussi être utilisées de manière plus traditionnelle comme preuve devant un tribunal pour démontrer que 2 saisies sont composées des mêmes produits illicites.

Les **conséquences** lors de la **mise en évidence de délits liés au trafic de stupéfiants** peuvent être **particulièrement lourdes**, par exemple : le fait de diriger ou d'organiser un groupement ayant pour objet la production, la fabrication, l'importation, l'exportation, le transport, la détention, l'offre, la cession, l'acquisition ou l'emploi illicites de stupéfiants est puni de la **réclusion criminelle à perpétuité** et de 7 500 000 euros d'amende (article 222-34 du Code pénal).

Durant cette étape, on obtient ce qu'on appelle un **profil** ou une **signature** qui peut être :

- chimique si les caractéristiques des composés sont uniquement d'origine chimique (impuretés organiques, inorganiques, solvants résiduels, produits de coupage ...). On obtiendra alors un profil d'impuretés, de solvants, d'adultérants, de diluants, etc.
- physique si ce sont des attributs plutôt macroscopiques comme l'apparence (couleur, forme, présence d'un logo, type de logo, etc.), ou l'emballage (type, présence d'adhésifs, etc.). Sur les emballages, des profils chimiques peuvent aussi être déterminés.
- des combinaisons des différents types de profil sont possibles.

Les niveaux d'information diffèrent suivant les types de profils recherchés.

L'héroïne, la cocaïne, le cannabis et les amphétamines de type stimulant (ATS en anglais, comprend l'amphétamine, la méthamphétamine et la 3,4-méthylènedioxy-N-méthylamphétamine – MDMA ou ecstasy) sont les stupéfiants les plus largement diffusés partout dans le monde et faisant l'objet d'un vaste trafic organisé. Ils constituent également ce qui est perçu comme un problème de santé majeur dans la plupart des pays consommateurs. Par conséquent, du point de vue du profilage, l'accent est mis principalement sur ces quatre substances illicites.

Le schéma ci-après, extrait du chapitre sur le profilage des drogues de l'encyclopédie des sciences forensiques ¹⁶, fournit le concept global du procédé de profilage des drogues (figure 7).

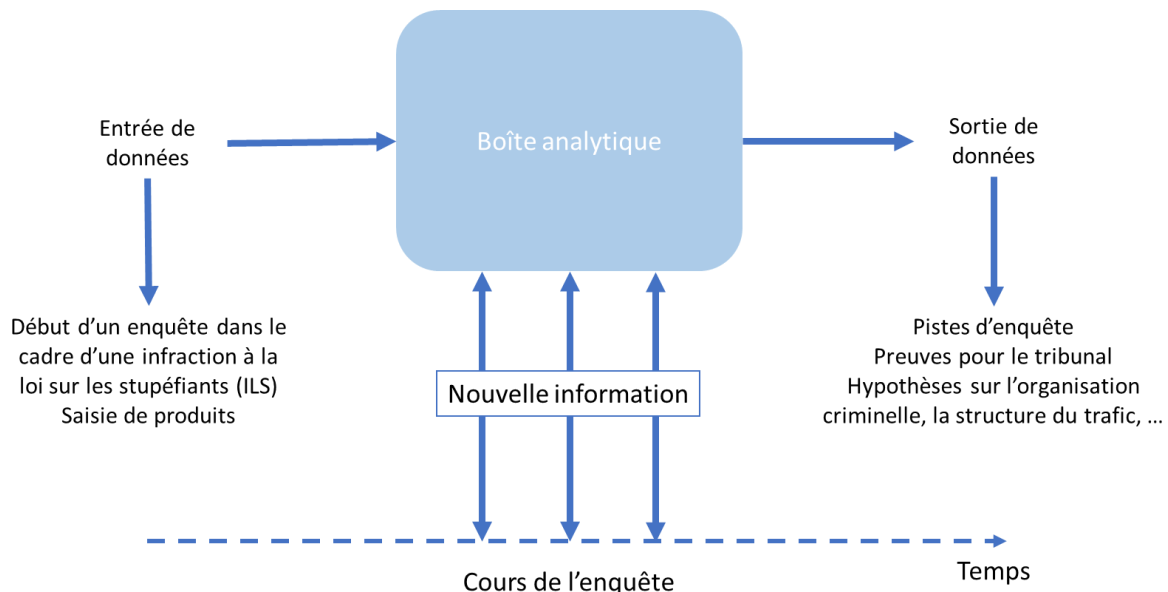


Figure 7 : concept global du procédé de profilage des drogues

Le point de départ est l'enquête dans le cadre d'une infraction à la loi sur les stupéfiants avec une saisie de produits. Des informations circonstancielles comme la date, le lieu, la provenance, la destination de la saisie sont alors disponibles partiellement ou en totalité. L'autre source d'informations provient de l'analyse de la saisie et de son interprétation. Les informations combinées peuvent être examinées et

interprétées dans la boîte analytique afin d'extraire des résultats pertinents pour la poursuite de l'enquête.

Tout au long de ce processus, l'enquête reste active et peut fournir de nouvelles informations ou utiliser celles qui sont apportées par l'analyse.

L'objectif principal est d'extraire des connaissances utiles pour mener des enquêtes complexes, comprendre les éléments clés d'une activité criminelle, formuler des hypothèses raisonnables sur les organisations criminelles ou sur la structure du trafic et enfin fournir des preuves au tribunal.

Au niveau de la boîte analytique, une approche itérative est suivie de la façon suivante :

- Constitution de connaissances tacites acquises par la recherche pour l'établissement de protocoles analytiques, la détermination de facteurs discriminants, d'indicateurs de la chaîne de production de la drogue, de transformation en produit final saisi dans la rue, la validation de ces paramètres à partir de cas antérieurs (stabilité dans le temps). Élément clé pour pouvoir extraire de la connaissance lors de nouveaux cas.
- L'action préliminaire pour les nouveaux cas consiste à collecter des éléments d'information obtenus à partir du trafic de drogue connu ou perçu. Cette étape est d'une importance majeure ; en effet, sans information, le processus de renseignement est arrêté et ne peut être réalisé. Ces éléments d'information sont la saisie de drogue elle-même et toutes les informations (appel téléphonique, données de l'informateur, transaction bancaire, etc.) recueillies par différents enquêteurs.
- Les données d'analyses (c.-à-d. le profil établi) et les autres données sont ensuite organisées dans une base regroupant les informations pertinentes.
- À partir de là, le processus d'interprétation est enclenché pour produire des renseignements « forensique » (informations extraites des profils de saisie), qui sont combinés à des renseignements d'enquête. Ces sources de renseignements combinées produisent des connaissances utiles sur le trafic de drogue à la fois dans une perspective opérationnelle, mais aussi stratégique. Ce processus d'interprétation est appelé **analyse criminelle**.

2.4. Échantillonnage de stupéfiants

2.4.1. Échantillonnage pour des analyses d'identification (qualitatives)

Lors d'une saisie de stupéfiants, deux cas de figure peuvent se présenter, soit la totalité de la saisie est envoyée au laboratoire et à ce moment-là, si sa taille est trop importante, un échantillonnage sera réalisé, soit, une partie de la saisie est transmise. Il n'est alors pas possible de contrôler parfaitement la sélection des spécimens. Néanmoins, le travail du laboratoire est aussi de fournir des lignes directrices sur la façon de collecter des échantillons.

Quelles que soient les étapes d'analyses (identification, dosage et comparaison), pour fournir des résultats fiables et interprétables, la maîtrise de l'échantillonnage est primordiale.

La question de l'échantillonnage a été une des premières thématiques sélectionnées par le groupe de travail sur les drogues de l'ENFSI. Dès 2001, un sous-comité a été créé

afin de réaliser un manuel sur l'échantillonnage pour des analyses d'identification (appelées également analyses qualitatives). Une première version du manuel a vu le jour en 2003. En 2009, il a été repris par l'UNODC ¹⁷.

En 2012, une version révisée a été produite incluant un calculateur pour l'échantillonnage sous la forme d'un tableur Excel® ¹⁸.

Tous les documents sont accessibles en libre consultation sur le site de l'ENFSI : <https://enfsi.eu/about-enfsi/structure/working-groups/documents-page/documents/forensic-guidelines/>

Quand on parle **d'échantillonnage au sens général**, c'est qu'en principe, on se retrouve dans le cas où l'on ne peut pas réaliser des mesures sur tous les objets (ou unités supposées discrètes et homogènes) à disposition, soit parce qu'ils sont trop nombreux et cela engendrerait une énorme perte de temps, soit parce que la mesure est destructive ou l'inspection de tous les objets peut exposer la personne à un risque pour sa santé.

Ainsi, un échantillon représente une partie de cet ensemble appelé alors « population ».

On cherche alors à répondre à une question dans laquelle le terme « combien » intervient. On peut aussi retrouver le terme de « représentatif ». D'ordinaire, derrière tout cela, il y a la nécessité de prendre une décision sous incertitude, à savoir : « on aimerait connaître la quantité d'unités à prendre pour être sûr d'être représentatif de la population ou pour obtenir des résultats représentant fidèlement la population ». C'est le plus souvent sous cette forme que la question est posée : « **quelle taille d'échantillon faut-il prélever ?** », quel que soit le domaine dans lequel la personne travaille.

Par exemple : dans l'industrie, il peut s'agir de pièces fabriquées en grande quantité (comme des boulons ou des produits agroalimentaires, pharmaceutiques, etc.) sur lesquelles on va rechercher la présence de défauts ou non. « Combien de pièces doit-on prélever ? ».

Dans le cas d'une enquête, il pourrait y avoir des articles individuels tels que des sacs ou des dispositifs de stockage de données électroniques. Chaque article peut contenir ou non quelque chose d'illégal (soit de la drogue dans le cas d'un trafic de drogue ou des images pornographiques dans un autre cas). « Combien d'objets doit-on analyser ? ».

Bien évidemment, il n'est pas possible de répondre à cette question directement. Une imprécision ou incertitude est introduite lorsque l'inférence est faite de l'échantillon à la population, car l'ensemble de la population n'est pas inspecté. Cependant, cette incertitude peut être quantifiée de manière probabiliste. Il est alors nécessaire de définir quel est le degré d'incertitude qu'on est prêt à accepter (ou le risque qu'on est prêt à prendre) avant de prendre la décision.

La **question**, « quelle taille d'échantillon faut-il prélever ? » peut être **formulée plus précisément** de la façon suivante : « on cherche à être certain à $100 \times p\%$ que $100 \times \theta_0\%$ ou plus de l'ensemble (la population) contient une substance illégale ou est défectueux lorsque toutes les unités échantillonnées contiennent une substance illégale ou sont défectueuses. Si on considère par exemple, $p = 0.95$ et $\theta_0 = 0.5$, la question devient : « quelle taille d'échantillon faut-il prendre pour dire qu'il y a 95% de chances pour que la proportion d'unités dans la saisie (ou le lot ou la population) contenant une substance illégale (ou défectueuse) soit supérieure à 50% ? ».

Par **représentativité**, on peut entendre que les conclusions sont généralisables, c'est-à-dire que les conclusions d'une étude s'appliquent à d'autres situations que celle des cas de l'étude.

Comment obtenir un échantillon représentatif ? Ce n'est pas simple et sujet à incertitude¹⁹. C'est pourquoi certains principes d'échantillonnage doivent être suivis pour que l'échantillon soit représentatif de la population. C'est donc la méthode qui donnera aux échantillons leur qualité de représentativité²⁰.

La statistique a montré (et démontré mathématiquement) qu'un échantillon représente très fidèlement la population lorsqu'il est tiré de manière aléatoire. Il existe plusieurs façons de sélectionner des échantillons de façon aléatoire. La méthode la plus simple pour obtenir un échantillon aléatoire consiste à tirer au sort, à l'aveugle, un certain nombre d'éléments de la population, après les avoir bien mélangés (par exemple dans un sac ou une urne). Un tel échantillon porte le nom d'**échantillon aléatoire simple**.

Trois types d'échantillonnage peuvent être envisagés. Il s'agit de l'échantillonnage aléatoire simple, de l'échantillonnage stratifié ou à plusieurs degrés et de l'échantillonnage composite²¹. Dans le cas de **l'échantillonnage de drogues**, on s'est focalisé sur le premier.

Deux approches statistiques

Pour déterminer la taille de l'échantillon, 2 approches statistiques peuvent être envisagées, l'une fréquentiste et l'autre bayésienne. Leur application à l'échantillonnage de substances illicites est décrite ci-après.

L'approche fréquentiste

Dans cette approche, l'hypothèse sous-jacente est qu'une **proportion fixe, mais inconnue dans la saisie** contient une drogue illicite. La proportion calculée dans un échantillon permet d'estimer la proportion dans la saisie. La proportion de l'échantillon variera toutefois d'un échantillon à l'autre. Si l'on tire un grand nombre d'échantillons, la plupart donneront des résultats très proches de la valeur réelle, quelques-uns s'écarteront un peu plus de celle-ci et un petit nombre donnera des résultats plus éloignés encore. **Ainsi, en tirant un seul échantillon (en fait, on n'en a qu'un dès le départ), on peut partir du principe suivant : la probabilité que le résultat de l'estimation soit très proche de la valeur réelle est très grande ; inversement, il y a peu de chance qu'il s'écarte grandement de cette valeur. Par conséquent, il est possible de déterminer de combien le résultat pourrait s'écarter de la valeur réelle.**

Les méthodes fréquentistes fournissent une confiance $(1-\alpha)100\%$ (par exemple 95% si α est choisi égal à 0,05), qu'avec une proportion dans l'échantillon donnée, la proportion dans la saisie est d'au moins $k100\%$ (par exemple 90% si k est choisi égal à 0,9). Cela signifie que si la proportion échantillonnée est effectivement celle supposée, on aurait raison de dire qu'il s'agit d'une saisie contenant au moins 90 % de drogue dans 95 cas sur 100.

Les distributions de probabilités utilisées dans cette approche sont la loi hypergéométrique et la loi binomiale. Dans l'une, on suppose que les échantillons sont prélevés sans remise (la taille de la population diminue à chaque sélection, effet négligeable si population de grande taille). La distribution hypergéométrique est discrète et tous les nombres appliqués dans les calculs doivent être des entiers non négatifs. Cette application est adaptée aux situations dans lesquelles le calcul de la taille de l'échantillon est basé sur un nombre seuil (K) ou une proportion seuil (k)

d'éléments positifs (drogue) attendus dans la population (saisie), comme spécifié par le laboratoire.

L'autre méthode est l'utilisation de la loi binomiale dans laquelle il y a remise. Elle est plus simple, mais ne peut être utilisée que dans des cas particuliers. En effet, cela signifie qu'une unité est replacée après avoir été échantillonnée et analysée avant que l'unité suivante ne soit échantillonnée. Bien sûr, cela est rarement pratiqué dans l'échantillonnage des drogues puisque les analyses sont destructives.

Néanmoins, lorsque la taille de la population est très grande (au moins 50 unités, préférablement plus) et que la taille de l'échantillon est relativement petite (comparée à la taille de la population), la loi hypergéométrique peut être approximée par la loi binomiale.

Dans l'approche **fréquentiste**, le calcul de la taille de l'échantillon est basé sur des **tests d'hypothèses** : la taille minimale d'un échantillon (n) est calculée en testant l'hypothèse nulle selon laquelle le nombre de positifs N_i dans la population est inférieur à K contre l'hypothèse alternative que le nombre de positifs N_i est au moins égal à K .

Le tableau ci-dessous fournit la taille d'échantillon requise pour garantir avec 95% ou 99% de confiance que la saisie contient au moins une proportion de k substances illicites si l'on s'attend à ce que 0, 1 ou 2 unités échantillonnées n'en contiennent pas (0, 1 ou 2 négatifs), dans le cas des grandes saisies.

Population de taille N	95 % de confiance			99% de confiance		
	k = 0.5	k = 0.75	k = 0.9	k = 0.5	k = 0.75	k = 0.9
0 négatif	5	11	29	7	17	44
1 négatif	8	18	46	11	24	64
2 négatifs	11	23	61	14	31	81

Tableau 1 : Pour de grandes saisies, tailles d'échantillons requises pour garantir avec un certain pourcentage de confiance (95 ou 99) que la saisie contient au moins une proportion k de drogues ($k = 50, 75$ et 90%) si l'on s'attend à ce que l'échantillon contienne 0, 1 ou 2 unités ne contenant aucune drogue.

L'approche bayésienne

Dans celle-ci, l'hypothèse sous-jacente est que la **proportion dans l'échantillon est connue et fixe**. Cette proportion est utilisée pour calculer des probabilités sur certaines valeurs de la **proportion dans la saisie inconnue**, qui à ce moment-là est encore **supposée variable**. Avec cette approche, il est possible d'intégrer certaines connaissances sur la saisie. La proportion dans la saisie n'est pas connue, mais l'analyste en a souvent une idée. Par exemple, si tous les plants d'une pépinière de cannabis se ressemblent, ils sont probablement tous des plants de cannabis. Il est également possible que l'analyste (s'il est débutant par exemple) n'ait aucune idée de la proportion et du type de drogue dans une saisie (par exemple des comprimés blancs sans logo). Ces diverses connaissances de départ se traduiront par différents modèles mathématiques pour estimer une taille d'échantillon souhaitée dans l'approche bayésienne.

Dans cette approche, une distinction peut aussi être faite entre l'échantillonnage avec remise et l'échantillonnage sans remise. Encore une fois, l'échantillonnage avec remise est plus simple et peut être utilisé comme approximation dans des situations où la taille de la population est d'au moins 50 et l'échantillon est relativement petit. Ici, la surestimation n'est pas un problème aussi important qu'avec la distribution binomiale.

C'est pourquoi l'échantillonnage avec l'approximation avec remise est beaucoup plus utilisé dans l'approche bayésienne.

Les bayésiens supposent que, bien que la proportion de la population ne soit pas connue, il peut y avoir une certaine idée de cette proportion qu'on appelle un a priori (c.-à-d. connaissance préalable). Cet a priori est représenté par une distribution de probabilité, dite distribution a priori de la proportion. Cette connaissance incertaine est combinée avec les informations fournies par l'échantillon à une distribution dite a posteriori des proportions, compte tenu des résultats de l'échantillon. Avec cette distribution a posteriori, il est possible de calculer directement la probabilité que la proportion de drogues soit au moins k (compte tenu des résultats de l'échantillon) sans utiliser de tests d'hypothèses ou d'intervalles de confiance. En effet, les bayésiens calculent directement $P(\theta > k \mid x, n)$ au lieu de $P(X > x \mid \theta > k, n)$ comme le font les fréquentistes.

Si une population est grande ($N > 50$) et que l'échantillon est relativement petit par rapport à celle-ci, la fonction de densité de probabilités pour la **proportion θ de positifs**, étant donné qu'un échantillon de taille n contient x positifs est une **distribution Bêta** de paramètres $x + a$ et $n - x + b$. a et b doivent être sélectionnés au préalable en fonction des connaissances a priori ou des hypothèses sur θ . Les connaissances préalables combinées aux informations sur les données (échantillon de taille n et nombre de positifs x dans l'échantillon) forment la **distribution a posteriori**. Comme dans les méthodes fréquentistes, on suppose à l'avance quel sera le nombre de positifs dans l'échantillon en question, et on adapte les conclusions si par la suite ce nombre n'est pas correct. Encore une fois, dans la plupart des cas, aucun négatif n'est attendu.

Outre le nombre attendu de positifs dans l'échantillon, il faut choisir une **distribution a priori**. En général, il s'agit également d'une **distribution Bêta**.

Une suggestion consiste à prendre les deux paramètres a et b égaux à 1, si on n'a aucune idée au préalable sur le contenu des unités. La distribution a priori est alors égale à une distribution uniforme. Une autre suggestion est de prendre a et b tous deux égaux à $\frac{1}{2}$ s'il y a un a priori selon lequel soit toutes les unités contiennent de la drogue, soit aucune unité n'en contient.

Si on choisit $b = 1$, et $a = 3$ (ou même plus) cela signifie qu'il existe une croyance préalable, basée sur l'inspection visuelle de la saisie, l'expérience ou autre, que probablement tout est de la drogue. Par exemple, 100 petits sachets en plastique refermables avec 3 liserés rouges (type minigrip®) de poudre blanche sont trouvés, tous emballés de la même manière, tous ayant le même poids, et tous sentant la cocaïne (odeur caractéristique).

Les tableaux 2a, 2b et 2c fournissent la taille d'échantillon requise pour garantir avec une probabilité de 95% ou 99% que la saisie contient au moins une proportion de k substances illicites si l'on s'attend à ce que 0, 1 ou 2 unités échantillonnées n'en contiennent pas (0, 1 ou 2 négatifs). La taille de la saisie est supposée grande ($N \geq 50$).

Dans le tableau 2a, on utilise les paramètres $a=1$ et $b=1$ dans le calcul des distributions Bêta si aucune information préalable n'est connue.

Dans le tableau 2b, on utilise $a=0.5$ et $b=0.5$ s'il est raisonnable de supposer que, soit tout est de la drogue (tout est positif), soit rien n'est de la drogue (tout est négatif).

Dans le tableau 2c, on utilise les paramètres $a=3$ et $b=1$ (ou des valeurs plus extrêmes), s'il y a des raisons de penser que la totalité ou la majeure partie de la saisie contient des drogues.

Paramètres de la distribution a priori Bêta $a = 1$ et $b = 1$ <i>Pas d'information connue</i>	95 % de confiance			99% de confiance		
	$k = 0.5$	$k = 0.75$	$k = 0.9$	$k = 0.5$	$k = 0.75$	$k = 0.9$
0 négatif	4	10	28	6	16	43
1 négatif	7	17	47	10	24	64
2 négatifs	10	23	63	13	31	83

Tableau 2a : Pour de grandes saisies. Taille d'échantillon requise pour garantir avec une probabilité de 95% ou 99% que la saisie contient au moins une proportion de k substances illicites si l'on s'attend à ce que 0, 1 ou 2 unités échantillonnées n'en contiennent pas (0, 1 ou 2 négatifs) lorsqu'on n'a aucun a priori ($a = 1$ et $b = 1$).

Paramètres de la distribution a priori Bêta $a = 0.5$ et $b = 0.5$ <i>Tout + ou tout -</i>	95 % de confiance			99% de confiance		
	$k = 0.5$	$k = 0.75$	$k = 0.9$	$k = 0.5$	$k = 0.75$	$k = 0.9$
0 négatif	2	8	26	4	14	41
1 négatif	5	15	45	8	22	62
2 négatifs	8	21	61	11	29	81

Tableau 2b : Pour de grandes saisies. Taille d'échantillon requise pour garantir avec une probabilité de 95% ou 99% que la saisie contient au moins une proportion de k substances illicites si l'on s'attend à ce que 0, 1 ou 2 unités échantillonnées n'en contiennent pas (0, 1 ou 2 négatifs) lorsqu'on a un a priori selon lequel tout est de la drogue ou rien n'est de la drogue.

Paramètres de la distribution a priori Bêta $a = 3$ et $b = 1$ <i>Tout +</i>	95 % de confiance			99% de confiance		
	$k = 0.5$	$k = 0.75$	$k = 0.9$	$k = 0.5$	$k = 0.75$	$k = 0.9$
0 négatif	3	7	18	5	12	32
1 négatif	7	15	39	9	21	55
2 négatifs	10	21	56	12	28	75

Tableau 2c : Pour de grandes saisies. Taille d'échantillon requise pour garantir avec une probabilité de 95% ou 99% que la saisie contient au moins une proportion de k substances illicites si l'on s'attend à ce que 0, 1 ou 2 unités échantillonnées n'en contiennent pas (0, 1 ou 2 négatifs) lorsqu'on a un a priori selon lequel tout est de la drogue.

La connaissance a priori va donc jouer un rôle important dans la diminution de la taille de l'échantillon, pour une probabilité de 99%, on passe de 43 à 32 suivant que l'on ne sait rien ou que l'on connaît bien le domaine. Ce nombre représente du temps de gagné et de l'économie (en solvant, instrument analytique, etc.).

Si la saisie est petite ($N < 50$) alors il est préférable de considérer le nombre de positifs dans les unités non examinées plutôt que la proportion de positifs. La fonction de densité de probabilités pour le nombre de positifs dans les unités non examinées Y , étant donné qu'un échantillon de taille n contient x positifs est une Bêta-Binomiale. Là encore, dans la plupart des cas, on s'attend à aucun résultat négatif.

Lignes directrices proposées aux laboratoires de police scientifique

Dans le manuel, l'ENFSI prodigue des conseils à plusieurs niveaux : national et international.

Au **niveau national/régional**, l'échantillonnage est une stratégie et il dépend fortement de l'objet à échantillonner, de la question posée et des objectifs. Les lois nationales et les pratiques juridiques dictent la plupart de ces éléments. Dans la pratique, il existe souvent une certaine liberté, ce qui signifie que les forces de police régionales, les tribunaux et les laboratoires peuvent avoir la possibilité de développer leur propre stratégie. L'échantillonnage doit être adapté à leurs besoins, satisfaisant pour le requérant, facile à comprendre. Il doit également tenir compte de la charge de travail des laboratoires, et être rentable.

La décision est donc laissée aux spécialistes du domaine de l'analyse des drogues, qui, en collaboration avec la direction de leur laboratoire choisiront et développeront une stratégie appropriée, satisfaisante pour leurs clients (police, tribunaux) et en accord avec eux.

Au **niveau international**, certaines saisies importantes présentent des aspects internationaux, signifiant que des suspects seront trouvés dans un ou plusieurs pays. Il a été jugé nécessaire d'avoir une stratégie raisonnable pouvant être largement soutenue par (la plupart) des chimistes forensiques dans les pays de l'UE et être utilisée comme ligne directrice par la police et les autorités compétentes.

L'ENFSI a donc proposé les recommandations suivantes :

- dois avoir une base facile à expliquer en termes de statistiques ;
- dois être facile à comprendre et être pratique, notamment à l'usage des policiers et des douaniers ;
- dois être réaliste et ne pas entraîner une augmentation de la charge de travail pour les laboratoires (ce qui se traduirait par des délais d'analyses inacceptables). Cela signifie une approche minimale et non maximale ;
- dois encore être raisonnablement défendable devant un tribunal.

C'est dans cette perspective que le laboratoire de Lyon a mis en place une stratégie d'échantillonnage suivant ces recommandations. **Des tailles d'échantillons, en fonction des cas les plus fréquents, sur la base de l'approche bayésienne ont été calculées et des formations aux personnels des laboratoires ont été dispensées.**

2.4.2. Échantillonnage pour des analyses quantitatives (dosages)

L'analyse quantitative d'une saisie de drogue sert à répondre à la question « quelle est la concentration d'une drogue dans une saisie ? ».

Il n'est donc pas possible d'utiliser la même approche statistique d'échantillonnage que celle recommandée pour l'analyse qualitative, car la quantité mesurée d'une substance illicite dans une saisie peut être fortement influencée par l'hétérogénéité de cette saisie. En outre, l'approche statistique de l'échantillonnage qualitatif s'il était appliqué à l'échantillonnage quantitatif ne serait pas réaliste, car il pourrait générer trop d'unités pour l'analyse quantitative. Il a donc fallu concevoir un protocole d'échantillonnage différent pour l'analyse quantitative.

En 2007, un sous-comité du WG-Drugs de l'ENFSI a été créé afin d'élaborer des directives d'échantillonnage appropriées pour l'analyse quantitative de différentes drogues illicites. Ces directives devaient être pratiques et économiques, tout en respectant les théories d'échantillonnage scientifiquement reconnues.

Ce sous-comité était composé de membres de Grande-Bretagne, d'Irlande du Nord, de Hongrie, de Suisse et de France. La présidente du sous-comité de 2007 à 2013 était,

Mme Anne Franc de Forensic Equity – Grande-Bretagne. En 2011, j'ai pris la présidence.

En 2010, un appel à projets financé par la Direction générale des affaires intérieures de la Commission Européenne dans le cadre du programme "Prévenir et combattre la criminalité" a été lancé. Un consortium a été créé et un **projet sur le sujet de l'échantillonnage de stupéfiants pour des analyses quantitatives** a été proposé. Il a été retenu et financé pour une durée de 3 ans (grant agreement number HOME/2010/ISEC/MO/4000001759/MP6 ; budget d'environ 30 000€ dédiés à des réunions de travail et à la publication du manuel). Les membres du projet étaient :

- Mr. Michael Bovens de l'institut des sciences forensiques de Zürich – Suisse,
- Dr. Julia Nagy et Dr. Tamas Csesztregi de l'institut hongrois des sciences forensiques, Budapest – Hongrie,
- Mme Anne Franc de Forensic Equity, Wokingham – Grande-Bretagne,
- Dr. Laurence Dujourdy de l'Institut National de Police Scientifique, Service Central des Laboratoires, Ecully - France.

La figure 8 résume la mise en place de la réflexion autour de cette thématique.

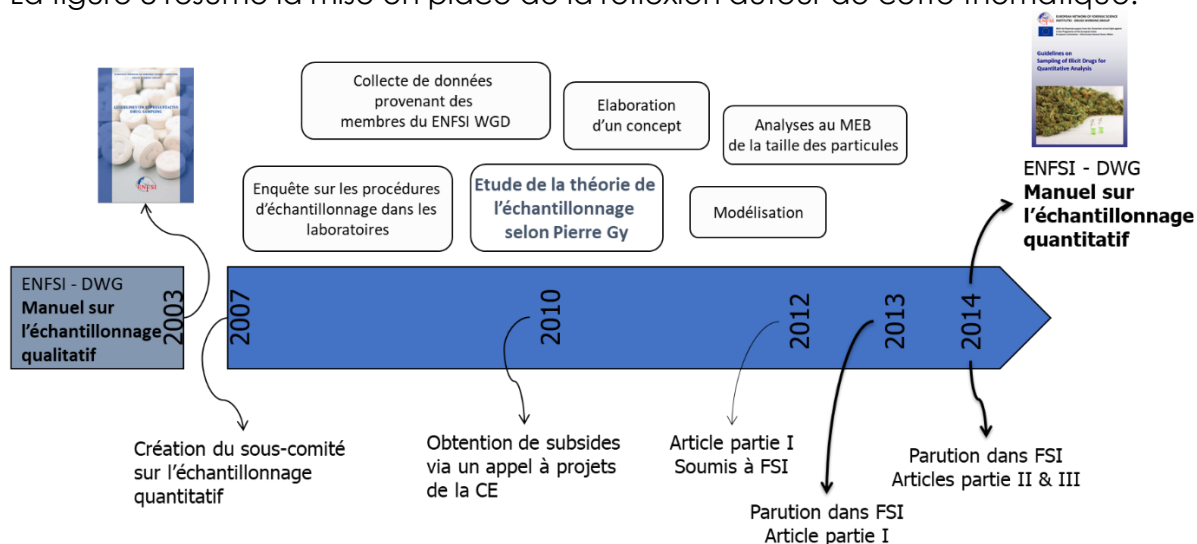


Figure 8 : chronologie de la réflexion sur l'échantillonnage quantitatif et production scientifique.

La réponse à la question « quelle est la concentration d'une drogue dans une saisie ? » est du type : $\bar{X} \pm U\%$ avec \bar{X} la moyenne de la concentration et U l'incertitude sur le calcul de celle-ci.

Le résultat est dépendant de 2 paramètres : l'homogénéité de la matière et la taille des particules dans la matière.

Dans le cas d'un échantillonnage qualitatif, le plan correspond à un nombre d'unités testées individuellement. Dans le cas d'un échantillonnage quantitatif, le plan correspond à l'élaboration d'un échantillon primaire composé d'unités d'une certaine quantité de matière à partir duquel, après un procédé d'homogénéisation, un spécimen est pesé pour être analysé afin de déterminer la concentration du composé d'intérêt.

Dans certains pays, la pureté de la drogue illicite lors d'une saisie peut affecter la condamnation pénale et, par conséquent, des résultats fiables pour une analyse quantitative sont cruciaux.

Les substances investiguées étaient les principaux stupéfiants rencontrés sur le marché international de la drogue, à savoir : l'héroïne, la cocaïne, l'amphétamine, les comprimés de MDMA et le cannabis (herbe et résine).

L'objectif du projet était de produire des recommandations pour des plans d'échantillonnage pratiques en vue de l'analyse quantitative.

Le but fondamental de l'échantillonnage pour l'analyse quantitative est de maintenir la concentration moyenne de la drogue dans le matériau depuis son état d'origine jusqu'au spécimen analytique.

Dans la pratique, lorsqu'il s'agit de saisies importantes, il n'est pas possible d'homogénéiser l'ensemble du matériau et les techniques d'analyse modernes ne nécessitent qu'une petite quantité (quelques milligrammes) de matière. Par conséquent, il est nécessaire d'effectuer des mesures sur une plus petite portion prélevée dans la masse, tout en maintenant la concentration moyenne de la substance illicite de la saisie initiale.

L'échantillonnage et l'analyse peuvent être considérés comme 2 processus distincts, chacun avec sa propre incertitude. Lorsqu'un laboratoire décide de démontrer la performance globale de l'ensemble du processus de quantification, dans le cadre de la validation d'une méthode d'analyse, il doit estimer l'incertitude totale provenant à la fois du processus analytique et du processus d'échantillonnage. La seule exception, lorsque l'échantillonnage n'a pas d'influence, est lorsque la totalité de la saisie est utilisée comme spécimen pour analyse.

La probabilité d'obtenir une seule portion du matériel en vrac représentant parfaitement la concentration de celui d'origine est extrêmement faible. Si plusieurs portions (incrément) sont prises, leurs caractéristiques varieront en raison de l'hétérogénéité des composés du matériau, mais la moyenne des concentrations des portions individuelles fournira une meilleure estimation de la concentration de tout le matériau.

L'un des principaux objectifs de notre étude était d'aligner l'incertitude de l'échantillonnage (exprimée en tant qu'écart-type relatif d'échantillonnage ou « RSD_{sampling} ») sur l'incertitude de la méthode analytique appliquée (exprimée en tant qu'écart-type relatif d'analyse ou « RSD_{analysis} »), qui est généralement de l'ordre de 2 à 5 % (pour des méthodes chromatographiques, principales techniques utilisées pour l'analyse des drogues).

Le schéma suivant explique le processus d'échantillonnage et d'analyse quantitative ²² (figure 9) :

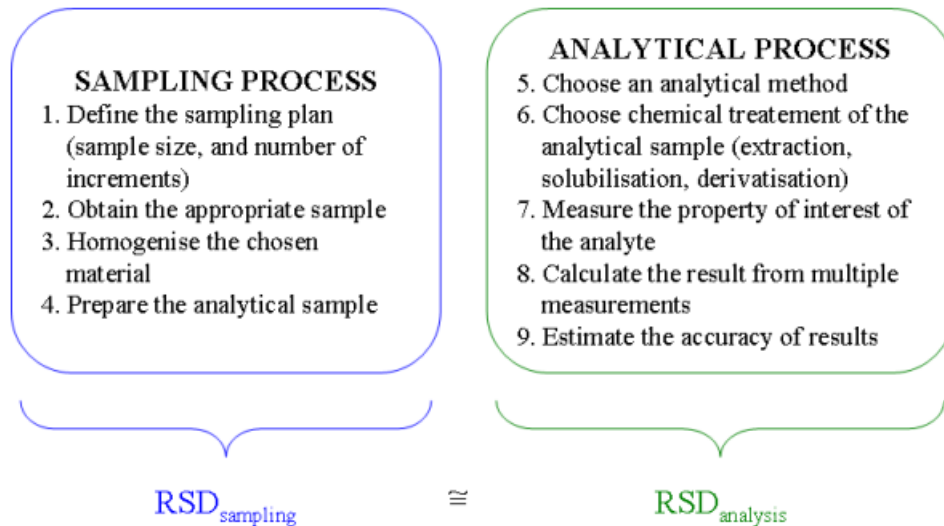


Figure 9 : Aperçu du processus d'échantillonnage et d'analyse quantitative.

Des études préliminaires avaient montré que les 2 facteurs clés ayant un effet important sur l'obtention d'un échantillon analytique représentatif d'une saisie en vrac étaient l'hétérogénéité de la poudre de drogue et la taille des particules de ses composants. On s'est donc focalisé dans un premier temps sur l'**hétérogénéité** des drogues. Les effets de la **taille des particules** ont été abordés dans un second temps. Dans un troisième temps, la **description d'une procédure d'échantillonnage** adaptée à l'analyse quantitative des drogues illicites basée sur les résultats théoriques et expérimentaux des 2 parties précédentes a été proposée.

Les 3 phases ont fait l'objet d'articles scientifiques publiés dans la revue spécialisée : « Forensic Science International (FSI) ».

La théorie de Pierre Gy^{23,24} a servi de base à l'élaboration de ces articles et du manuel sur l'échantillonnage quantitatif de drogues²². Il s'agit d'une théorie développée en référence à des problèmes miniers, mais qui rencontre de nombreuses applications dans d'autres domaines. Son objectif est de prédire la précision d'un échantillon donné pour représenter un lot donné. Les 2 hypothèses sous-jacentes sont :

- l'échantillon est probabiliste et sans biais,
- le lot peut être représenté par 2 phases distinctes :
 - o phase comprenant le minéral (ou la variable) d'intérêt
 - o phase formée du reste (gangue)

a) Première phase d'étude : l'hétérogénéité des drogues

Tous les matériaux réels sont hétérogènes. La question est de savoir comment et dans quelle mesure ?

Il existe deux types d'hétérogénéités :

- l'**hétérogénéité de constitution** : fait référence aux différences dans la constitution du matériau, c'est-à-dire à quel point les particules ou molécules individuelles sont similaires ou différentes. Elle caractérise la variabilité entre les fragments (ou particules) individuels (pour les solides) ou entre les molécules individuelles (pour les liquides et les gaz). La **figure 10** montre un exemple d'hétérogénéité de constitution pour de l'herbe de cannabis (matière la plus hétérogène),

- l'**hétérogénéité de distribution** : fait référence aux différences dans la façon dont les morceaux (fragments, particules ou molécules) sont distribués dans l'espace, c'est-à-dire à quel point le matériau est bien mélangé ou séparé en raison de la densité, de la taille des particules ou d'autres facteurs. La **figure 11** montre un exemple d'hétérogénéité de distribution dans le cas de l'herbe de cannabis.

L'hétérogénéité du matériau a une forte influence sur l'échantillonnage. Différents morceaux de matière contribuent différemment à la concentration moyenne (c'est-à-dire que la concentration du principe actif est différente dans les bourgeons, les feuilles et les tiges de cannabis). De plus, la répartition des différentes particules n'est pas uniforme dans les matériaux bruts (avant broyage).

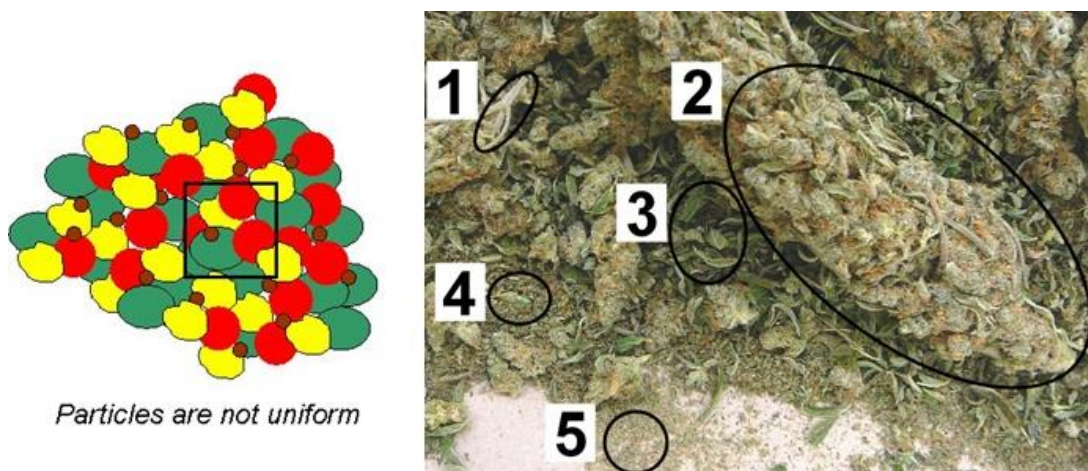


Figure 10 : Illustration de l'hétérogénéité de constitution d'un échantillon d'herbe de cannabis : (1) = tiges, (2) = bourgeons, (3) = grandes parties de feuilles, (4) = petites parties de feuilles et (5) = poudre.

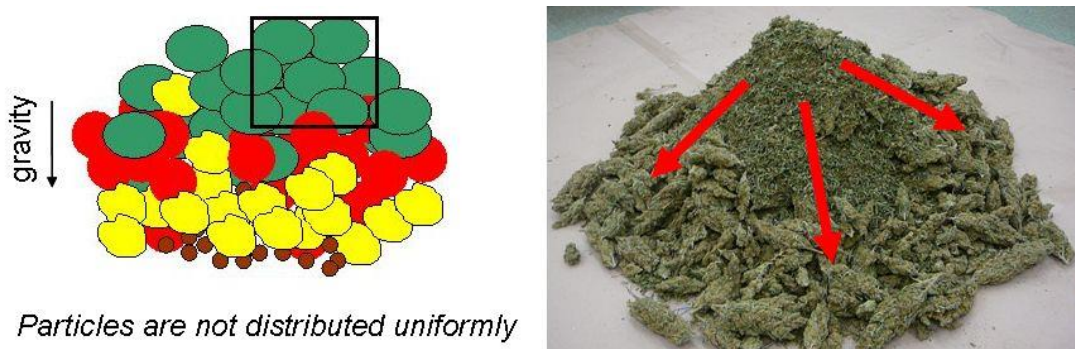


Figure 11 : Illustration de l'hétérogénéité de distribution avec l'échantillon d'herbe de cannabis de la figure 10.

L'hétérogénéité des principaux stupéfiants a été étudiée et des distributions d'hétérogénéité empiriques ont été établies pour pouvoir calculer les erreurs générées lorsqu'on calcule la concentration en principe actif. Un nombre d'incrémentes représentant une certaine quantité en grammes a été estimé pour chaque type de stupéfiant. ²⁵.

Une méthode générale pour obtenir un échantillon primaire représentatif du matériel saisi a été recommandée. Le schéma de la figure 12 montre la différence entre un échantillonnage arbitraire et un échantillonnage incrémental et l'erreur qui peut être commise en raison de la présence d'hétérogénéité. L'échantillon incrémental m_1 correspond au soi-disant « échantillon correct » comme expliqué dans la théorie de Pierre Gy ²³.

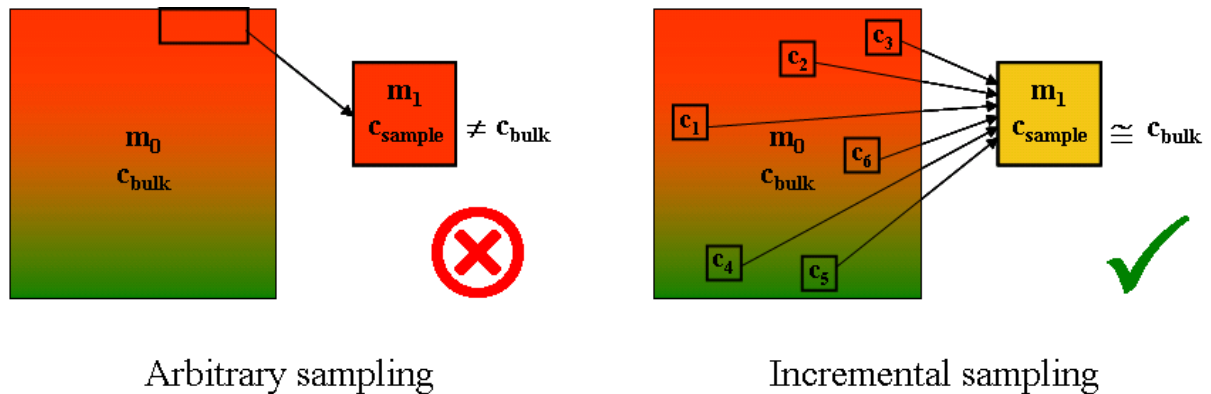


Figure 12 : Comparaison entre un échantillonnage arbitraire et un échantillonnage incrémental.

b) Deuxième phase d'étude : effet de la taille des particules

La partie II de cette série d'articles traitait des effets de la taille des particules au cours de la chaîne de réduction de masse et du processus de préparation de l'échantillon : de « l'échantillon primaire » à « l'échantillon analytique », où l'effet de la taille des particules est le paramètre le plus critique. Cette étude était basée sur la mesure de la taille des particules de différents matériaux de drogues illicites dans leur état d'origine et après homogénéisation à l'aide de procédures manuelles ou mécaniques.

Comme on l'a vu, l'échantillonnage est un processus générateur d'erreurs. L'hétérogénéité (sous ces 2 formes, hétérogénéité de constitution (CH) et hétérogénéité de distribution (DH)) est la principale source de toutes les erreurs d'échantillonnage. L'erreur fondamentale (FE) est l'erreur minimale générée lorsqu'on prélève un échantillon d'une masse donnée. Elle est influencée par la taille des particules.

L'erreur fondamentale (FE pour « fundamental error ») est inhérente à la distribution de la composition du matériau de l'échantillon et ne peut être modifiée sans changer la nature de l'échantillon. Cette erreur peut être réduite, soit en augmentant la masse de l'échantillon, soit en diminuant la taille des plus grosses particules qu'il contient, par broyage.

On s'est donc focalisé sur le calcul de FE, avec en particulier l'étude de l'influence de la taille des particules. L'objectif final étant d'obtenir un spécimen analytique pesant seulement quelques milligrammes et représentatif de l'échantillon global d'origine.

La relation entre l'écart-type relatif et la masse de l'incrément analytique pour des stupéfiants pulvérulents a été modélisée pour différentes puretés (de faible (< 10 %) à supérieure (≥ 100%)) et pour différentes granulométries nominales après broyage (de 300 μm à 30 μm).

En règle générale, plus la pureté du stupéfiant est élevée, plus la masse de l'incrément analytique requise est faible (comme on s'y attendait).

Calcul de la variance relative de l'erreur fondamentale FE

En substituant la taille de particule nominale mesurée dans les matériaux bruts (non broyés) dans l'équation d'échantillonnage (éq. 1), il a été démontré que le choix d'incrément de 1 g pour les échantillons primaires de stupéfiants en poudre et de résine de cannabis était correct. Pour le cannabis végétal, la taille appropriée d'incrément était plus grande (c'est-à-dire environ 1 bourgeon, ce qui équivaut à environ 1 à 5 g).

$$s_{FE}^2 = \frac{0.125cd^3}{m_1}$$

Équation 1

Avec :

d : diamètre de la plus grosse particule. Il a été déduit des résultats de cette étude.

c : facteur minéralogique, défini comme l'hétérogénéité maximale générée par le constituant (analyte) d'intérêt dans le lot.

m_1 : masse de l'incrément (g) (m_0 étant la masse du lot (g))

En substituant les valeurs de taille de particules nominales mesurées dans le matériau échantillonné dans l'éq. 1, il a été démontré que la quantité appropriée de substance illicite nécessaire à la préparation d'un échantillon analytique pouvait être considérablement réduite, tout en maintenant un RSD de 2 à 5 %, ce qui était conforme au RSD pour le reste du processus analytique. Cela s'appliquait aux matériaux de pureté « normale » c'est-à-dire entre 1 % et 100%, pas en traces. Plus la taille nominale des particules est petite, plus la masse de matériau nécessaire pour un échantillon analytique est faible ²⁶.

c) Troisième phase d'étude : plan d'échantillonnage général

La troisième partie de la série d'articles sur la base des résultats des études précédentes proposait un plan d'échantillonnage général (représenté sous la forme d'un organigramme). Un tableau a été construit montrant l'incertitude d'échantillonnage attendue lorsqu'un nombre particulier d'incrément aléatoires est prélevé et combiné pour former un seul échantillon primaire. Il comprenait également une masse ou nombre recommandé(e) d'incrément, de 1 g pour les drogues en poudre et pour la résine de cannabis, 1 comprimé pour la MDMA et 1 bourgeon pour l'herbe de cannabis sous forme de bourgeon (type I). Il s'est avéré que d'autres formes plus hétérogènes de cannabis (type II) nécessitaient d'autres méthodes d'échantillonnage plus traditionnelles (mortier + pilon).

En se référant au tableau ci-dessous (tableau 3), les laboratoires peuvent s'assurer que l'incertitude d'échantillonnage pour une saisie de drogue particulière peut être minimisée de façon à ce qu'elle se situe dans la même gamme que leur incertitude analytique pour cette drogue.

$S_{\text{écha.}} \text{ (RSD\%)} \\ \text{Pour un matériau avec hétérogénéité} \\ S_{\text{écha.}} = S_{\text{hétérogénéité}} / \sqrt{n}$			
Nombre d'incrément (n)	20% d'hétérogénéité (pour bourgeon de cannabis végétal, incrément d'1g trop petit)	12% d'hétérogénéité (comprimés)	10% d'hétérogénéité (poudres & résine de cannabis)
5	8.9%	5.4%	4.5%
10	6.3%	3.8%	3.2%
15	5.2%	3.1%	2.6%
20	4.5%	2.7%	2.2%
50	2.8%	1.7%	1.4%
100	2.0%	1.2%	1.0%
Taille (masse ou nombre) d'incrément recommandée	1 bourgeon	1 comprimé	1 gramme

Tableau 3 : formule et exemples de la manière dont l'écart-type relatif attendu peut être réduit en augmentant le nombre d'incrément.

Des exemples de cartes de contrôle de l'échantillonnage ont été également proposés (cf. un exemple figure 13). Elles peuvent être construites à partir de l'étendue ou de l'étendue relative (carte R) des résultats (x_1 et x_2) provenant d'au moins deux répétitions complètes du processus d'échantillonnage sur une cible d'échantillonnage. L'étendue relative peut être utilisée pour des échantillonnages réalisés sur des cibles différentes. Elles peuvent aussi être construites à partir de l'écart-type (carte S).

Une carte de contrôle est utile pour contrôler l'écart-type du processus d'échantillonnage, en utilisant la ligne centrale comme étant l'écart-type relatif du processus d'échantillonnage (s_{rel}) estimé comme étant la moyenne des écarts types relatifs d'échantillonnages répétés sur différentes cibles et les limites d'alerte et d'action, calculées en fonction du nombre de résultats individuels par cible d'échantillonnage.

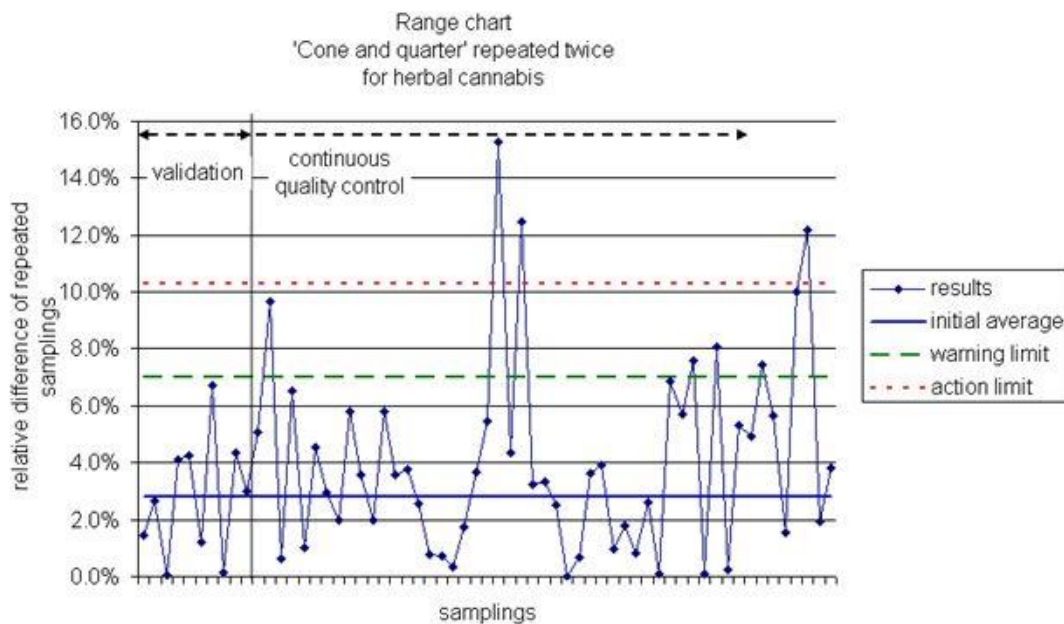


Figure 13 : Carte de contrôle sur l'étendue, carte R.

Les éléments développés dans cette série d'articles ont servi pour établir un deuxième manuel de recommandations pour les laboratoires qui ont besoin de réaliser de l'échantillonnage de drogues pour des analyses quantitatives ²². Il s'agissait du principal livrable du projet financé par la CE et l'ENFSI.

Des ateliers de formation ont également été proposés lors de la 21^e réunion annuelle du groupe de travail sur les drogues du réseau ENFSI et de la 7^e conférence européenne de l'académie des sciences forensiques en 2015 (références 18 et 19 du CV page 9).

2.5. Validation de méthodes analytiques

Si l'on reprend le procédé global des analyses de stupéfiants, l'échantillonnage représentait la phase préliminaire aux analyses pour les 3 étapes : identification, dosage et profilage (figure 14).

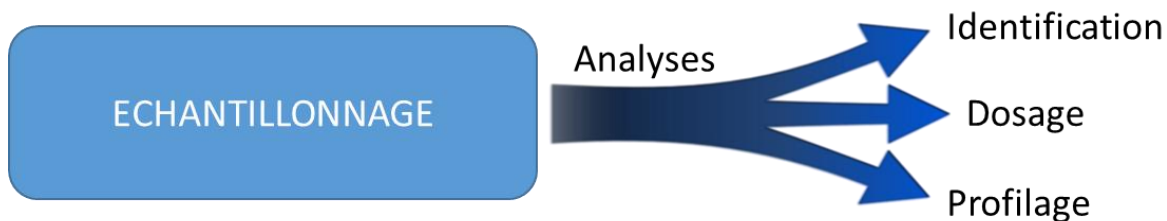


Figure 14 : Échantillonnage = phase préliminaire à l'analyse.

On va s'intéresser maintenant à la validation de méthodes analytiques dans le cadre de la démarche ISO 17025 pour l'accréditation des laboratoires d'essais et d'étalonnages dans lesquels s'inscrivent les laboratoires de police scientifique. Il est à noter que l'échantillonnage fait partie des exigences de cette norme.

2.5.1. Cadre de la norme ISO 17025

La norme impose aux laboratoires de valider les méthodes d'analyses qu'ils utilisent. Cela inclut les méthodes standard, non standard et développées en laboratoire.

Ce sont les analyses d'identification et de dosage qui ont été concernées en priorité et je me suis intéressée plus particulièrement à la validation des méthodes de dosage. Les analyses de profilage, quant à elles, n'ont pas fait l'objet d'une validation dans le cadre de cette norme.

La problématique était la suivante : plusieurs méthodes avaient été développées pour, d'une part, réduire les temps d'analyse (< 15 min) et, d'autre part, pallier les problèmes d'approvisionnement et de coût en hélium (gaz vecteur qui était utilisé auparavant sur les appareils de chromatographie gazeuse). Ces méthodes devaient être validées en vue d'une accréditation ISO 17025.

Pour cela, plusieurs critères doivent être vérifiés ²⁷ :

- La **sélectivité** : capacité de la méthode à établir de manière univoque l'existence de la substance à analyser en présence d'autres composés potentiellement présents.
- La **fonction de réponse** : elle traduit, à l'intérieur de l'intervalle de dosage, la relation existant entre la réponse (signal) et la concentration (quantité) en substance présente dans l'échantillon (caractérisé par sa masse incrémentale m_i définie dans le paragraphe précédent). La fonction de réponse monotone la plus simple qui exprime cette relation est appelée « courbe d'étalonnage ».
- La **linéarité** : capacité à l'intérieur d'un certain intervalle de dosage à obtenir des résultats directement proportionnels à la concentration (quantité) en substance à examiner dans l'échantillon. L'exigence de linéarité s'applique aux résultats c.-à-d. concentration calculée = $f(\text{conc. Introduite})$ et non pas aux réponses c.-à-d. signal = $f(\text{conc. Introduite})$.
- La **justesse** : exprime l'étroitesse de l'accord entre la valeur moyenne obtenue à partir d'une série de résultats d'essai et une valeur qui est acceptée soit comme une valeur conventionnellement vraie, soit comme une valeur de référence acceptée. La justesse fournit une indication sur les erreurs systématiques.
- La **fidélité** : exprime l'étroitesse de l'accord (degré de dispersion, ou écart-type) entre une série de mesures provenant de multiples prises d'un même échantillon homogène dans des conditions prescrites. La fidélité fournit une indication sur les erreurs dues au hasard. Elle traduit uniquement la distribution des erreurs aléatoires et n'a aucune relation avec la valeur vraie ou spécifiée. Elle peut être évaluée à 3 niveaux : la **répétabilité**, la **fidélité intermédiaire** (intralaboratoire) et la **reproductibilité** (interlaboratoire).
 - **Répétabilité** : conditions dans lesquelles les résultats d'essai indépendants sont obtenus par la même méthode sur des échantillons d'essai identiques dans le même laboratoire, par le même analyste, utilisant le même équipement et pendant un court intervalle de temps,
 - **fidélité intermédiaire** : conditions dans lesquelles les résultats d'essai indépendants sont obtenus par la même méthode sur des échantillons d'essai identiques dans le même laboratoire, par différents analystes et utilisant des équipements différents et pendant un intervalle de temps donné
 - **reproductibilité** : conditions dans lesquelles les résultats d'essai sont obtenus par la même méthode sur des échantillons d'essai identiques

dans différents laboratoires, par différents analystes et utilisant des équipements différents.

- L'**exactitude** : exprime l'étroitesse de l'accord entre le résultat d'essai et la valeur de référence acceptée, aussi appelée « valeur conventionnellement vraie ». L'étroitesse de l'accord ainsi observée est la résultante de la somme des erreurs systématiques et aléatoires, en d'autres termes l'erreur totale liée au résultat. Par conséquent, l'exactitude est l'expression de la somme de la justesse et de la fidélité.
- Le **seuil de détection** (limit of detection : LOD) d'une méthode d'analyse est la plus petite quantité d'une substance à examiner dans un échantillon pouvant être détectée, mais non quantifiée comme une valeur exacte.
- Le **seuil de quantification** (limit of quantification : LOQ) est la plus petite quantité d'une substance à examiner dans un échantillon pouvant être quantifiée dans les conditions expérimentales décrites avec une fidélité et une exactitude définies. Cette définition s'applique au seuil de quantification inférieur, mais il n'est pas rare de rencontrer également un seuil de quantification supérieur.
- Le **domaine de travail** d'une méthode d'analyse est la gamme de concentration entre les niveaux de quantification inférieur et supérieur (ces valeurs incluses) pour laquelle il a été démontré que la méthode est appropriée quant à sa fidélité, son exactitude et sa linéarité, en utilisant la méthode décrite.

La vérification de ces critères s'est basée sur les recommandations de la commission de la Société Française des Sciences et Techniques Pharmaceutiques (SFSTP) formulées en 3 articles ²⁷⁻²⁹. Elles reposent sur l'utilisation du **profil d'exactitude**, basé sur la notion d'erreur totale (biais + écart type), permettant de simplifier l'approche de la validation d'une procédure analytique tout en contrôlant le risque associé à son utilisation.

Quand on valide une méthode d'analyse quantitative, on cherche à répondre à la question : « Comment déclarer, de la façon la plus adéquate possible, valide ou non valide une méthode analytique ? »

En général, une méthode analytique ne donne pas la quantification directe de l'analyte à doser. L'échantillon de concentration inconnue x est analysé dans un appareil qui renvoie un signal y . Pour obtenir la valeur x il est nécessaire de passer par une courbe d'étalonnage (figure 15) :

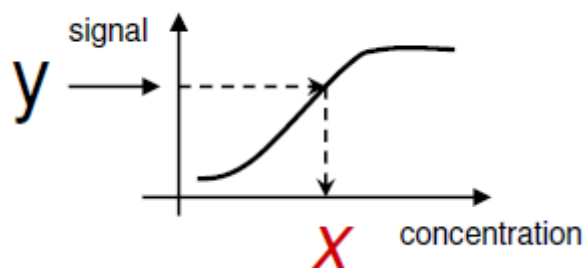


Figure 15 : Courbe d'étalonnage permettant d'obtenir la valeur de la concentration d'un analyte x en fonction du signal donné par l'appareil de mesure.

L'objectif est de permettre la détermination, le plus exactement possible, de chaque quantité inconnue que le laboratoire devra quantifier ; c'est-à-dire donner aux laboratoires ainsi qu'aux autorités, des garanties que chaque futur résultat obtenu en routine sera suffisamment proche de la vraie valeur inconnue de l'analyte dans l'échantillon. Ceci se traduit en termes de garanties par une proportion π de résultats obtenus supérieure à une proportion minimum de résultats inclus dans $\pm \lambda$; λ étant la limite d'acceptation prédéfinie, c'est-à-dire la valeur limite de l'écart toléré entre la valeur vraie (μ_T) (inconnue) et les résultats obtenus (μ étant la moyenne des résultats) (cf. figure 16). Il est à noter que ces limites peuvent être fixées par le client ou par une obligation réglementaire, parfois en fonction du niveau de concentration.

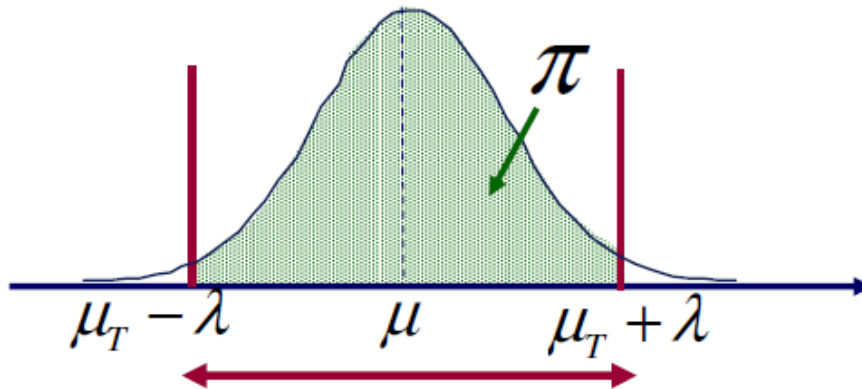


Figure 16 : Illustration d'un point de vue statistique de l'objectif de la validation d'une méthode quantitative.

2.5.2. Principe du profil d'exactitude

Dans cette optique, l'INPS avait signé un contrat d'engagement avec la société Arlenda en date du 27 juillet 2008 pour évaluer les performances analytiques des méthodes de dosage dans les domaines de la toxicologie médico-légale et de l'analyse de stupéfiants selon les protocoles de validations les plus récents ^{27,27,29,30} grâce à l'application web E-noval[®].

Elle permet d'estimer la justesse, la fidélité ainsi que l'exactitude d'une méthode analytique et aide également à évaluer l'effet potentiel de la matrice pour l'étalonnage. Elle génère des profils d'exactitude permettant la prise de décision pour le modèle d'étalonnage (calibration en anglais) et la validité de la méthode. Elle est basée sur le **concept de l'Erreur Totale** ³¹.

Le profil d'exactitude est réalisé par l'application directe des principes décrits dans les normes de la série ISO 5725 ³²⁻³⁵ et de la norme NF V 03-110 ³⁶. Un modèle statistique pour estimer l'exactitude (qui comprend la justesse et la fidélité) d'une méthode ou de résultats est proposé. Ce modèle décrit une mesure z d'une grandeur Z sous la forme (Eq. 2) :

$$Z = \bar{X} + B + E$$

Équation 2

où \bar{X} est la moyenne générale de l'échantillon homogène, B la composante du biais du laboratoire sous condition de répétabilité et E l'erreur aléatoire survenant dans chaque mesurage, sous condition de répétabilité.

Dans le cas d'une étude interlaboratoire, la composante du biais B provient des laboratoires, mais elle peut aussi provenir de toute autre source d'incertitude, comme le jour, l'opérateur, l'instrument... dans le cas d'une étude intralaboratoire.

Dans le vocabulaire statistique, un tel modèle est dit « modèle d'analyse de la variance à un seul facteur à effets aléatoires » ou aussi « modèle des composantes de la variance » (en anglais, One-Way ANOVA Random Effect Model). Ce modèle est très répandu dans divers domaines d'analyses (agronomie, pharmacie, environnement, etc.).

Outre les méthodes statistiques pour calculer les critères d'exactitude, les normes de la série ISO 5725 détaillent l'organisation de la collecte des données et les précautions à respecter pour éviter tout risque d'erreurs.

À partir de l'écart-type de fidélité intermédiaire ou de reproductibilité, calculé selon les recommandations de la norme ISO 5725-2³³, un intervalle de tolérance β contenant une proportion $\beta\%$ de futurs résultats est construit²⁸.

E-noval® produit également un rapport complet de validation avec les calculs statistiques nécessaires sans les expliciter. C'est une application basée sur le logiciel statistique SAS® pour le calcul de données. Il est accessible via une page web avec un droit de connexion. L'accès à une connexion est payant.

En raison du coût engendré par l'achat de connexions (une dizaine) à cette application pour les sections des laboratoires et du fait qu'elle peut être considérée comme une « boîte noire » (les modalités de calculs ne sont pas clairement explicitées, mais ils existent dans la littérature), **une solution de remplacement intitulée projet VALMETH a été envisagée**. Il s'agissait de développer en interne notre propre application statistique.

La démarche a été de programmer dans R (langage de programmation et logiciel libre destiné aux statistiques et à la science des données soutenu par la R Foundation for Statistical Computing³⁷) tous les calculs statistiques utilisés lors de la validation de méthodes de dosages et de générer les profils d'exactitude.

VALMETH devait reprendre la méthodologie statistique de la Commission SFSTP.

Les analystes ne connaissant généralement pas la programmation R, une interface web, hébergée sur un Réseau Général de Transport (RGT) du ministère de l'Intérieur, a été créée pour plus de facilité. Les données ne devaient pas être stockées ; seuls les rapports édités pouvaient être enregistrés par l'utilisateur sur le réseau de son laboratoire.

a) Base des calculs statistiques mis en œuvre pour la prise de décision

Usuellement, avec des résultats quantitatifs, les critères de décision portent sur les 2 types d'erreurs que sont l'erreur systématique ou justesse (biais = distance entre le résultat et la valeur vraie) et l'erreur aléatoire ou fidélité (variance = paramètre statistique chiffrant la dispersion des mesures autour de la moyenne).

Il faut ainsi montrer que :

- en moyenne, les résultats de la méthode sont comparables à ceux fournis par un laboratoire émetteur (justesse) (tests de performance interlaboratoire par exemple),
- la variabilité des résultats (fidélité) de la méthode est acceptable.

Suivant le guide élaboré par la SFSTP, 3 approches statistiques sont envisageables pour analyser les résultats quantitatifs de méthodes non normalisées.

- 1) La première approche statistique est purement descriptive : elle est basée sur l'estimation et la comparaison des moyennes et des coefficients de variation des résultats.

- 2) La seconde est l'approche par la différence, qui utilise des tests statistiques d'hypothèses bilatéraux comme le test t de Student.
- 3) La troisième est l'approche par l'équivalence, qui compare les paramètres des intervalles de confiance aux limites d'acceptation.

Par rapport à ces approches classiques, il en existe une originale qui est basée sur le profil d'exactitude, utilisant des intervalles de tolérance et l'erreur de mesure totale, comme outil de décision statistique simple.

L'approche descriptive (1) n'est pas un outil de décision adéquat, car elle ne contrôle ni le risque de rejeter une méthode valable ni celui d'accepter une méthode non valable en réalité.

L'approche par la différence (2) utilise le biais absolu, le biais relatif et le taux de recouvrement. Elle est basée sur l'utilisation d'un test bilatéral t de Student avec l'hypothèse nulle suivante :

$$H_0 : \text{biais} = 0 \Rightarrow H_0 : \text{biais relatif} = 0 \% \Rightarrow H_0 : \text{recouvrement} = 100 \%$$

avec $\text{biais} = x_i - \mu_v$, $\text{biais relatif} = [(x_i - \mu_v) / \mu_v] * 100$ et $\text{recouvrement} = (x_i / \mu_v) * 100$
 x_i = résultat obtenu avec la méthode d'analyse à valider
 μ_v = "vraie valeur" de l'échantillon (inconnue)

Une méthode est déclarée non biaisée et précise (qualifiée à tort) et donc valide en utilisant la règle de décision : quand les limites de confiance à 95 % du biais relatif moyen incluent la valeur de 0 %.

Mais ce test est inadéquat dans le cadre de la validation de méthodes. En effet, plus la variance est importante, c.-à-d. pire est la précision, plus l'intervalle de confiance contiendra la valeur 0% de biais relatif. Et plus la variance est faible, c.-à-d. meilleure est la précision, plus l'intervalle de confiance aura tendance à ne pas contenir la valeur de 0% conduisant à rejeter la méthode ce qui n'est pas l'objectif.

L'approche par l'équivalence (3) comme outil de décision par opposition à l'approche par la différence, utilise une limite de décision définie en accord avec les objectifs de la méthode à valider et permet d'éviter les contradictions décrites ci-devant.

Les hypothèses du test sont exprimées comme suit :

$$H_{01} : \text{biais relatif} \leq -\Delta \text{ contre } H_{11} : \text{biais relatif} > -\Delta$$
$$H_{02} : \text{biais relatif} \geq \Delta \text{ contre } H_{12} : \text{biais relatif} < \Delta,$$

où $\Delta(\%)$ est la limite de décision, c.-à-d. la différence maximale tolérée pour le biais relatif de la méthode

Les graphiques ci-après illustrent les décisions qui peuvent être faites suivant l'approche utilisée pour 5 situations différentes : figure (17a) approche par la différence et figure (17b) approche par l'équivalence en utilisant des limites de décision à +/- 5% représentée par les lignes rouges horizontales.

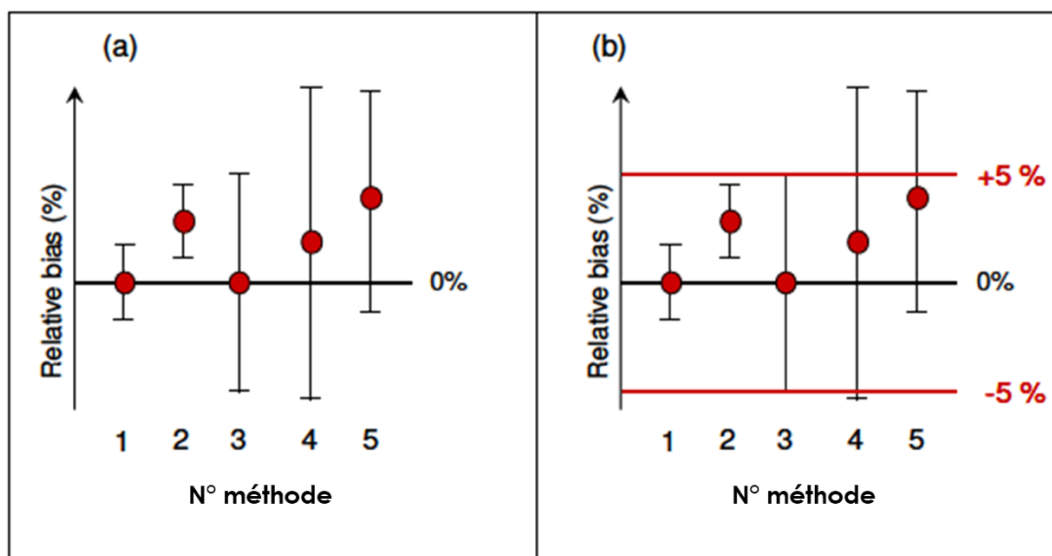


Figure 17 : Illustration de la différence de prise de décision suivant les deux approches statistiques utilisées : (a) approche par la différence, (b) approche par l'équivalence (d'après ³⁸).

Sur la figure (17a), la méthode 2 serait exclue avec l'approche par la différence puisque le 0% du biais relatif n'est pas compris dans l'intervalle à 95% alors que les méthodes 4 et 5 aux variances beaucoup plus élevées seraient quant à elles acceptées. Avec l'approche par l'équivalence montrée sur le graphique (17b), les méthodes 4 et 5 seraient rejetées et la 2 acceptée.

Cependant, l'approche par l'équivalence comme outil décisionnel n'est pas suffisante ; il ne répond pas à l'objectif de la validation de méthode. En effet, ce n'est pas parce que le biais relatif et la fidélité intermédiaire relative (Coefficient de Variation CV ou Relative Standard Deviation RSD en anglais) sont inclus dans des limites prédéfinies que les résultats sont acceptables (cf. figure 18). Néanmoins, l'opposé est toujours vrai : des résultats acceptables sont obtenus par des méthodes analytiques ayant un biais et une fidélité acceptables.

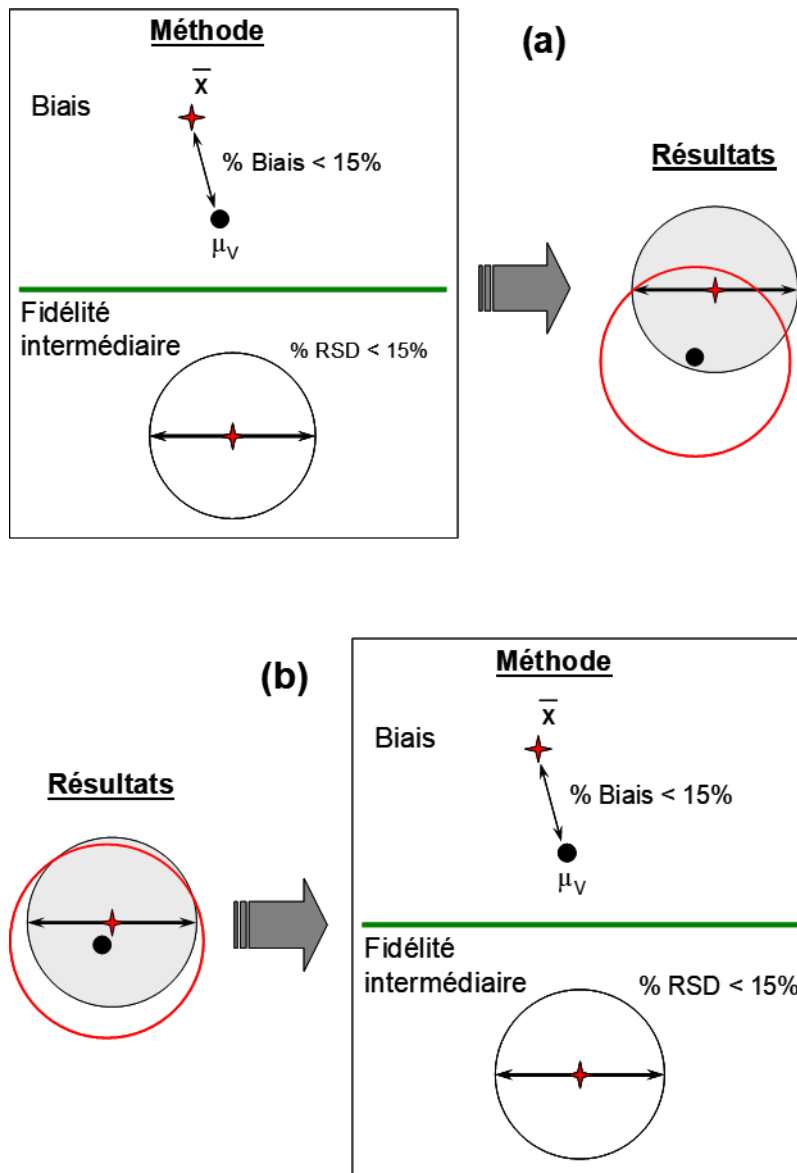


Figure 18 : Illustration qu'une bonne méthode ne donne pas forcément de bons résultats (a), mais de bons résultats sont obtenus avec une bonne méthode (b). μ_V est la valeur vraie, \bar{x} est la moyenne des résultats obtenus, le cercle rouge correspond aux limites d'acceptation des résultats $\pm 1\%$.

Dans ce contexte, l'utilisation du profil d'exactitude incluant des limites d'acceptation est l'outil de décision le plus adapté. Le profil d'exactitude construit sur des intervalles de tolérance permet de décider si une méthode analytique est capable de donner assez de résultats à l'intérieur des limites d'acceptation ou non (figure 19).

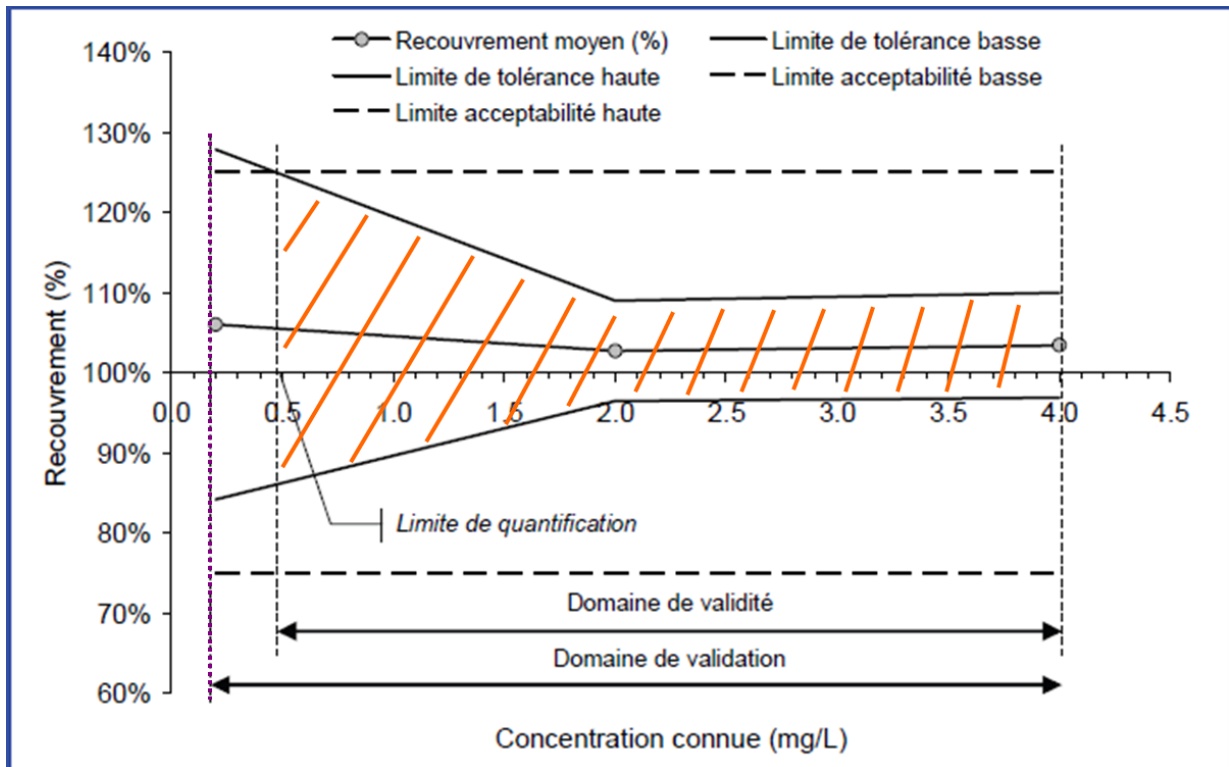


Figure 19 : Illustration du profil d'exactitude exprimé par le taux de recouvrement.

Sur la figure 19, chaque petit cercle gris représente la concentration moyenne retrouvée pour un niveau de concentration et quantifie la justesse (recouvrement moyen). Les lignes pointillées horizontales délimitent l'intervalle d'acceptabilité et les segments continus l'intervalle de tolérance calculé à partir des écarts-types de fidélité intermédiaire pour chaque niveau. Les traits verticaux délimitent le domaine de validité, dans lequel la méthode est capable de produire une proportion élevée et connue de résultats acceptables.

Pour utiliser le profil d'exactitude en vue de valider une méthode, il faut avoir fixé les **2 critères de décision** suivants :

- les **limites d'acceptabilité $\pm\lambda$** . Elles servent à traduire les objectifs pratiques des utilisateurs. Elles délimitent un intervalle autour de la valeur de référence. Le plus souvent, ces limites sont réglementaires ou issues de la réglementation. Mais dans le cas où il n'existe pas de référence établie, il convient de prendre en compte les attentes des utilisateurs finaux, comme une limite de quantification (LQ) donnée, par exemple, ou les valeurs issues des essais collaboratifs.
- La **proportion β** . Elle représente la proportion de futurs résultats qui seront en moyenne compris dans les intervalles de tolérance. La valeur choisie pour β dépend largement du champ d'application (contrôle sanitaire, contrôle de fabrication, etc.). Il est évident que plus la proportion est petite, par exemple 70 %, plus la méthode risque de produire des résultats qui ne correspondent pas aux spécifications annoncées.

Si l'analyste est prêt à prendre, par exemple un risque de 5%, il sera capable, à la fin de la validation de sa méthode, de garantir qu'en moyenne 95% (proportion β) des résultats futurs donnés par la méthode seront inclus dans les limites d'acceptabilité fixées en fonction des besoins (par exemple, 1% ou 2% pour l'analyse

environnementale, 5% pour les spécialités pharmaceutiques, 15% pour la bioanalyse, etc. ou les observations obtenues à partir de cartes de contrôles).

Nota bene

L'intervalle de tolérance est différent de l'intervalle de confiance. L'intervalle de confiance caractérise un paramètre statistique. Par exemple, l'intervalle de confiance de la moyenne permet d'établir avec un niveau de confiance donné un intervalle autour de la moyenne observée dans lequel la moyenne vraie a de grandes chances de se trouver. Il s'exprime sous la forme Moyenne $\pm k_{IC}$ (écart-type de la moyenne), où k_{IC} est le facteur de couverture de l'intervalle de confiance.

L'intervalle de tolérance caractérise d'un côté la dispersion des résultats déjà obtenus, mais permet aussi de prévoir ceux qui seront obtenus lors de futures mesures. Il s'exprime aussi sous la forme Moyenne $\pm k_{IT}$ (écart-type de fidélité), mais où k_{IT} est un autre facteur de couverture, calculé de façon complètement différente de k_{IC} ³⁹.

Autrement dit, un intervalle de tolérance est un intervalle incluant X% de la population, et ce, avec une probabilité p. La différence majeure est que l'intervalle de tolérance fait référence aux individus de la population (à la distribution de la population) alors que l'intervalle de confiance lui fait référence à un paramètre de la population. L'intervalle de tolérance sera donc beaucoup plus large que l'intervalle de confiance, parce que les observations sont souvent plus dispersées que la moyenne par exemple. Comme indiqué supra, les limites d'acceptabilité et la proportion β servant à calculer les intervalles de tolérance dépendent strictement du contexte d'utilisation de la méthode et doivent être adaptées à chaque cas.

D'une façon générale, pour être en cohérence avec les modèles statistiques utilisés, il faut disposer de répétitions effectuées sous condition de fidélité intermédiaire ou de reproductibilité sur des échantillons homogènes. Un protocole « minimal » est nécessaire à l'établissement du profil d'exactitude²⁷.

Règles de décision

Dès que l'intervalle de tolérance sort de l'intervalle d'acceptabilité, on peut conclure que la méthode n'est plus capable de fournir suffisamment de résultats acceptables, en fonction des choix faits au départ de l'étude.

Définition du domaine de validité

Le domaine de validité est déterminé par la zone du domaine de validation dans laquelle la méthode fournit une proportion de résultats acceptables au moins égale à β . Il est limité par une borne inférieure qui équivaut à la limite de quantification inférieure (LQi) et une borne supérieure qui correspond à la limite de quantification supérieure (LQs).

b) Mise en place de l'application VALMETH

Pour réaliser cette mise en place, en mars 2010, j'ai encadré un stagiaire, M. Ronan Maron, en 1^{re} année du master Santé et Population de l'université Claude Bernard – Lyon I (UCBL), dont la tutrice universitaire était la Docteur Anne-Béatrice Dufour, maître de conférences à l'UCBL, laboratoire de biométrie et biologie évolutive. Les résultats du stage ont permis d'obtenir une architecture pour le développement de l'application en lien avec le service informatique de l'Institut.

Une première version de VALMETH a été validée en mars 2012 et mise en circulation en avril 2012.

La figure 20 montre la page d'accueil de l'application avec la possibilité d'insérer automatiquement des données au format csv ou manuellement avec la possibilité d'utiliser 2 types de données, des réponses ou des résultats.

En effet, les instruments analytiques des laboratoires de l'Institut peuvent fournir soit des réponses (des courbes d'étalonnage seront ajustées et utilisées pour calculer les résultats estimés à partir des standards de validation), soit directement des résultats (un rapport de validation sans courbe d'étalonnage sera obtenu).

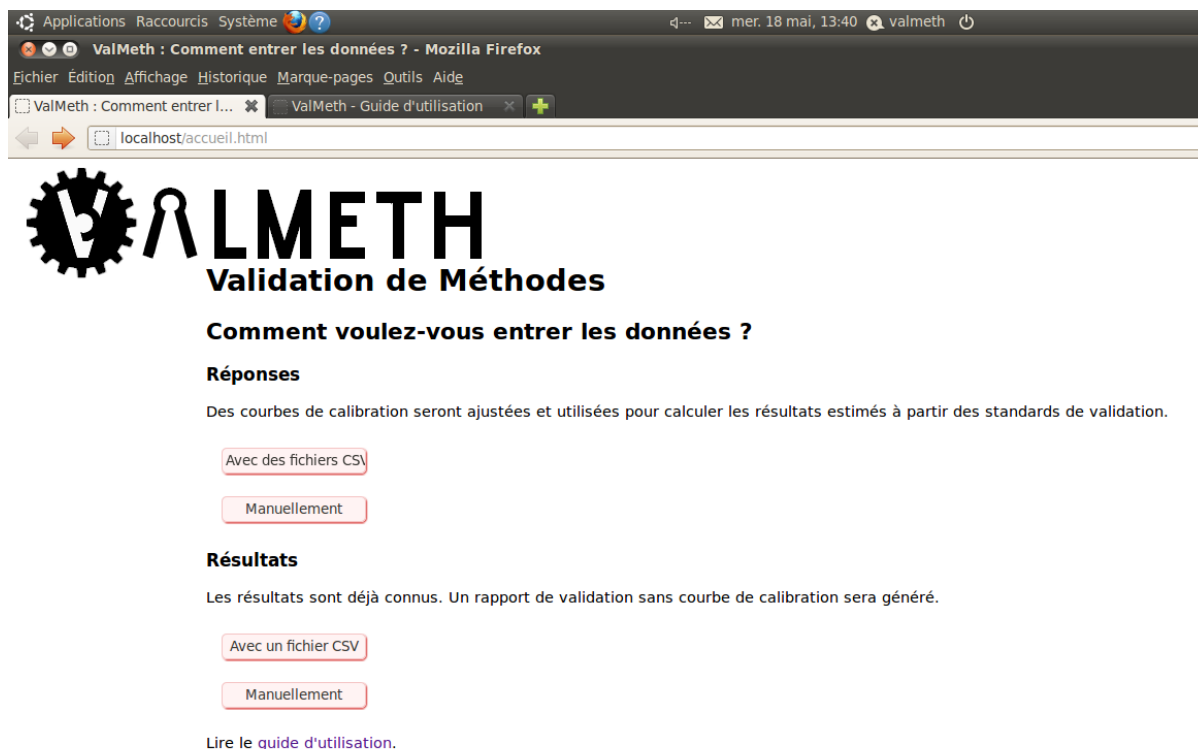


Figure 20 : Page d'accueil de l'application interne VALMETH pour la validation de méthodes analytiques quantitatives

Le programme propose à l'utilisateur de tester plusieurs modèles de régression pour déterminer la courbe d'étalonnage. Si plusieurs modèles sont sélectionnés, ils sont triés dans un tableau par ordre croissant d'indice d'exactitude :

1. régression linéaire passant par 0,
2. régression linéaire passant par 0 et utilisant un niveau choisi,
3. régression linéaire passant par 0 et utilisant le niveau maximum,
4. régression linéaire,
5. régression linéaire pondérée par $1/X$,
6. régression linéaire pondérée par $1/X^2$,
7. régression linéaire après transformation logarithmique de la concentration et de la réponse,
8. régression linéaire après transformation racine carrée de la concentration et de la réponse,
9. régression quadratique,
10. régression quadratique pondérée par $1/X$,
11. régression quadratique pondérée par $1/X^2$.

La figure 21 montre un exemple de données utilisées pour valider une méthode de dosage d'héroïne :



Implémentation sous R de statistiques utilisées pour la validation de méthodes quantitatives.

Application à des données en sciences forensiques.



Identifiant de l'échantion	Série	Niveau	Concentration (mg/ml)	Réponse
HER-S1-L1-SV01	1	1	0.04	0.032151957
HER-S1-L1-SV02	1	1	0.04	0.031836887
HER-S1-L1-SV03	1	1	0.04	0.032009571
HER-S1-L2-SV01	1	2	1.2	1.0510058
HER-S1-L2-SV02	1	2	1.2	1.062675842
HER-S1-L2-SV03	1	2	1.2	1.027696186
HER-S1-L3-SV01	1	3	3.2	2.791146857
HER-S1-L3-SV02	1	3	3.2	2.77884389
HER-S1-L3-SV03	1	3	3.2	2.766174671
HER-S2-L1-SV01	2	1	0.04	0.032447257
HER-S2-L1-SV02	2	1	0.04	0.032474284
HER-S2-L1-SV03	2	1	0.04	0.032450977
HER-S2-L2-SV01	2	2	1.2	1.0435548
HER-S2-L2-SV02	2	2	1.2	1.044112607
HER-S2-L2-SV03	2	2	1.2	1.042137479
HER-S2-L3-SV01	2	3	3.2	2.708815359
HER-S2-L3-SV02	2	3	3.2	2.729057381
HER-S2-L3-SV03	2	3	3.2	2.715921412
HER-S3-L1-SV01	3	1	0.04	0.030253421
HER-S3-L1-SV02	3	1	0.04	0.031436449
HER-S3-L1-SV03	3	1	0.04	0.031406888
HER-S3-L2-SV01	3	2	1.2	1.089424529
HER-S3-L2-SV02	3	2	1.2	1.060573068
HER-S3-L2-SV03	3	2	1.2	1.056570571
HER-S3-L3-SV01	3	3	3.2	2.779199863
HER-S3-L3-SV02	3	3	3.2	2.780914624
HER-S3-L3-SV03	3	3	3.2	2.7694853

TAB. 3 – Standards de validation.

Type	Équation	Paramètres	R
Droite passant par 0	$Y = \beta X$	β	$\ln(Y \sim X - 1)$
Droite	$Y = \alpha + \beta X$	α, β	$\ln(Y \sim X)$
Fonction quadratique	$Y = \alpha + \beta X + \gamma X^2$	α, β, γ	$\ln(Y \sim X + I(X \wedge 2))$

TAB. 4 – Fonction de réponses.

Figure 21 : Exemple de résultats d'analyses pour la validation de la méthode de dosage de l'héroïne.

Un tri des modèles est proposé en fonction de différents indices de performance. Il permet d'aider l'utilisateur à prendre une décision sur le modèle à appliquer. Il est à noter que ces indices (figure 22) ne font pas partie des calculs permettant de valider une méthode et ne sont donc qu'une aide à la décision.

Modèle	Limites inférieure et supérieure de quantification (Unité inconnue)	Indice d'exactitude	Indice d'intervalle de dosage	Indice de fidélité	Indice de justesse	Générer un rapport de prévalidation	Afficher le profil d'exactitude
Régression linéaire pondérée par $1/X^2$ (dans la matrice)	[0.04, 2.9]	0.9146	1.0000	0.7674	0.9969	Prévalidation	Profil d'exactitude
Régression linéaire pondérée par $1/X$ (dans la matrice)	[0.04, 2.9]	0.9145	1.0000	0.7680	0.9960	Prévalidation	Profil d'exactitude
Régression linéaire après transformation logarithmique (dans la matrice)	[0.04, 2.9]	0.9126	1.0000	0.7602	1.0000	Prévalidation	Profil d'exactitude

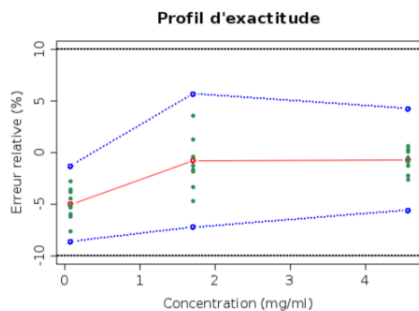
Figure 22 : Exemple d'indices calculés pour l'aide à la décision.

Ces indices sont décrits dans le guide d'utilisation de VALMETH.

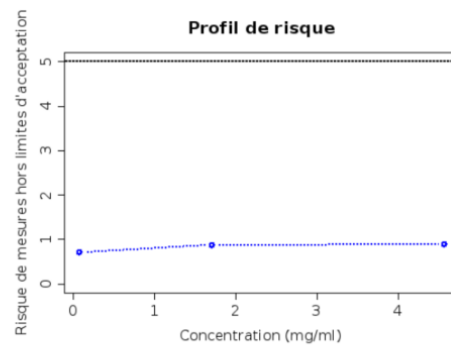
Le modèle de régression choisi par l'utilisateur est repris dans le texte du rapport. Un tableau récapitulant les paramètres de la régression est généré ainsi que des courbes d'étalonnage pour chaque série.

Un rapport avec tous les calculs et les graphiques est fourni à l'utilisateur. La figure 23 montre un extrait d'un rapport avec le profil d'exactitude et le profil de risque pour le dosage de l'héroïne.

Dans la matrice



La ligne rouge continue représente le biais, les lignes discontinues bleues définissent les limites de l'intervalle de tolérance attendues au niveau β et les lignes pointillées noires sont les limites d'acceptation. Les points sont quant à eux l'erreur relative des concentrations calculées et sont représentées en fonction de leurs concentrations cibles.



La ligne pointillée noire représente le niveau de risque maximum choisi, à savoir : 5 %.

Niveau de concentration	Moyenne des concentrations introduites (mg/ml)	Limites de tolérances β -expectation (mg/ml)	Limites de tolérances β -expectation relatives	Risque (%)
1	0.071	[0.06484, 0.07195]	[-8.67%, -1.34]	0.7037
2	1.710	[1.586, 1.807]	[-7.25%, 5.685]	0.8761
3	4.570	[4.313, 4.764]	[-5.625, 4.244]	0.8776

Exactitude de la méthode obtenue en considérant le modèle : régression linéaire pondérée par $1/X^2$

Figure 23 : Profil d'exactitude et profil de risque pour le dosage de l'héroïne extrait d'un rapport généré par VALMETH.

L'application Valmeth ne permet pas de calculer l'incertitude de mesure globale de la méthode. En effet, l'arrondi des résultats, la pureté des matériaux de référence et les incertitudes des appareils, par exemple, ne sont pas inclus. Cette incertitude est calculée postérieurement.

La collaboration avec le Dr Dufour s'est poursuivie sur le sujet et **a débouché sur 2 articles scientifiques**, l'un, plus général sur la statistique appliquée au dosage de la cocaïne paru dans les Annales Pharmaceutiques Françaises ⁴⁰ et l'autre, portant sur l'évaluation d'un modèle de régression pondérée pour les dosages de l'héroïne et de la cocaïne, paru Forensic Science International ⁴¹.

2.6. Profilage chimique des stupéfiants

La 3^e étape dans l'analyse des stupéfiants, telle que décrite dans le paragraphe 2.3.2. *Objectifs des analyses*, est la détermination de rapprochements entre saisies par l'établissement d'un profil chimique. À l'instar du profil génétique, le profil chimique d'une substance est hautement spécifique et peut indiquer sa filiation. En effet, la composition en éléments traceurs ne peut être reproduite d'un lot de fabrication à un autre. Il existe donc une signature chimique spécifique à chaque lot. Ainsi, lorsque 2 spécimens possèdent des profils similaires, on peut en déduire une origine commune, c'est-à-dire que les 2 saisies proviennent d'un même lot de fabrication, initialement plus grand, qui a été divisé au cours de la distribution ⁴².

Dans ce cadre de profilage de drogue, l'accent est mis sur le processus scientifique qui permet d'extraire des profils à partir de saisies de drogue. Une distinction majeure concerne la nature du matériau de départ. Si la matière première est « d'origine végétale » (cocaïne, héroïne et cannabis) ou provient de précurseurs synthétiques (ATS), le processus de fabrication est fortement affecté et influence le type de données pouvant être extraites des produits saisis comme l'indique la figure 24.

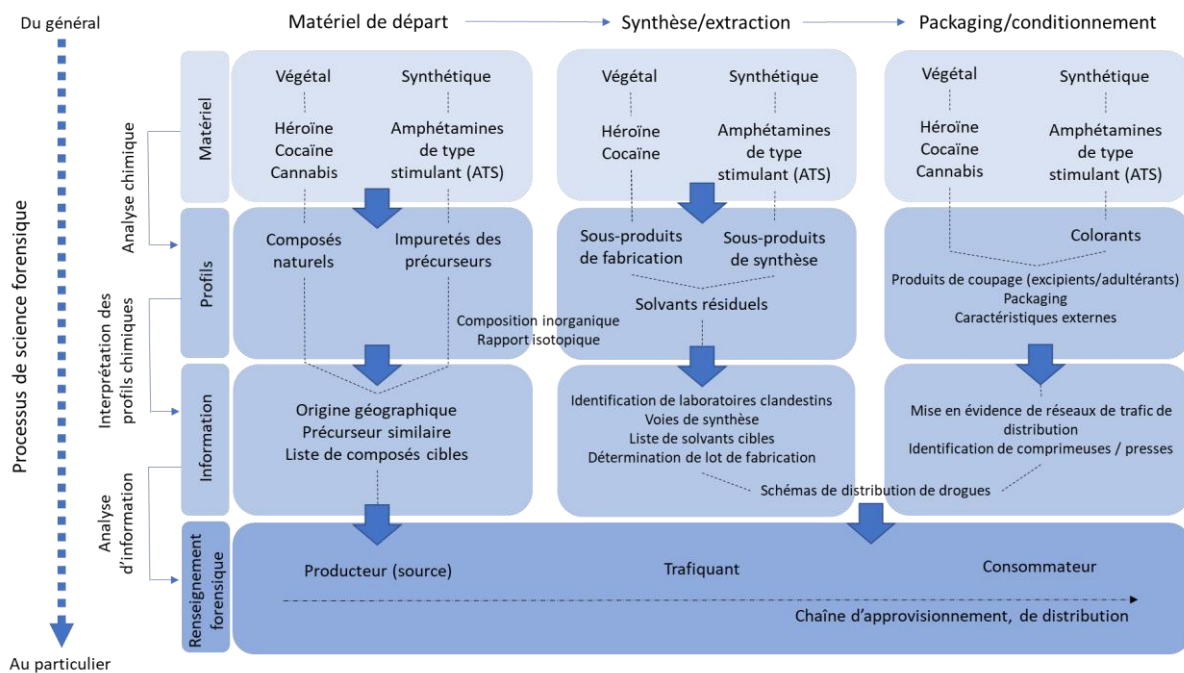


Figure 24 : Processus de science forensique dans le cas du profilage de drogues d'après Esseiva & Margot ¹⁶

Le profilage chimique exploite le fait que les **stupéfiants ne sont jamais purs**. En effet, les drogues saisies sont quasi systématiquement adultérées par des produits de coupage. Quand bien même elles ne le seraient pas, les trafiquants qui fabriquent ces drogues de façon illicite ne disposent pas des mêmes ressources nécessaires pour purifier leur production que l'industrie pharmaceutique par exemple. Dans le cas des drogues d'origine naturelle (telles que la cocaïne et l'héroïne), de nombreux alcaloïdes naturellement présents dans la plante, mais aussi des produits de dégradation et des solvants résiduels du procédé de fabrication, sont retrouvés dans le produit final. Il en va de même pour les drogues de synthèse, dans lesquelles des traces de produits secondaires de réaction, de catalyseur ou de réactif sont trouvées

dans le produit fini. Chacun de ces éléments (produit de coupage, alcaloïde naturel, impuretés de synthèse, produit de dégradation, solvants résiduels) peut constituer un paramètre sur la base duquel les échantillons sont comparés.

Cette comparaison peut servir des objectifs de renseignement (stratégique, tactique) ou avoir valeur de preuve.

L'établissement de tels rapprochements est d'un grand intérêt pour les services opérationnels qui peuvent ainsi combiner d'autres éléments d'enquête avec les résultats d'analyse et démanteler des réseaux de trafic de stupéfiants. Elle permet notamment de mettre en évidence un lien entre un revendeur et un consommateur, information pouvant s'avérer pertinente dans le cadre d'une garde à vue.

De plus, la nature et l'abondance des impuretés sont caractéristiques d'une voie de synthèse ou d'un procédé d'extraction. Il est par conséquent possible de déterminer quelle voie a été employée pour produire ledit lot. Cette seconde information est également intéressante pour les services d'enquête.

2.6.1. Processus de profilage

Voici une illustration (figure 25) des étapes du processus de profilage des stupéfiants et les méthodes analytiques de recherche de composés en fonction de l'information à fournir (d'après ⁴³).

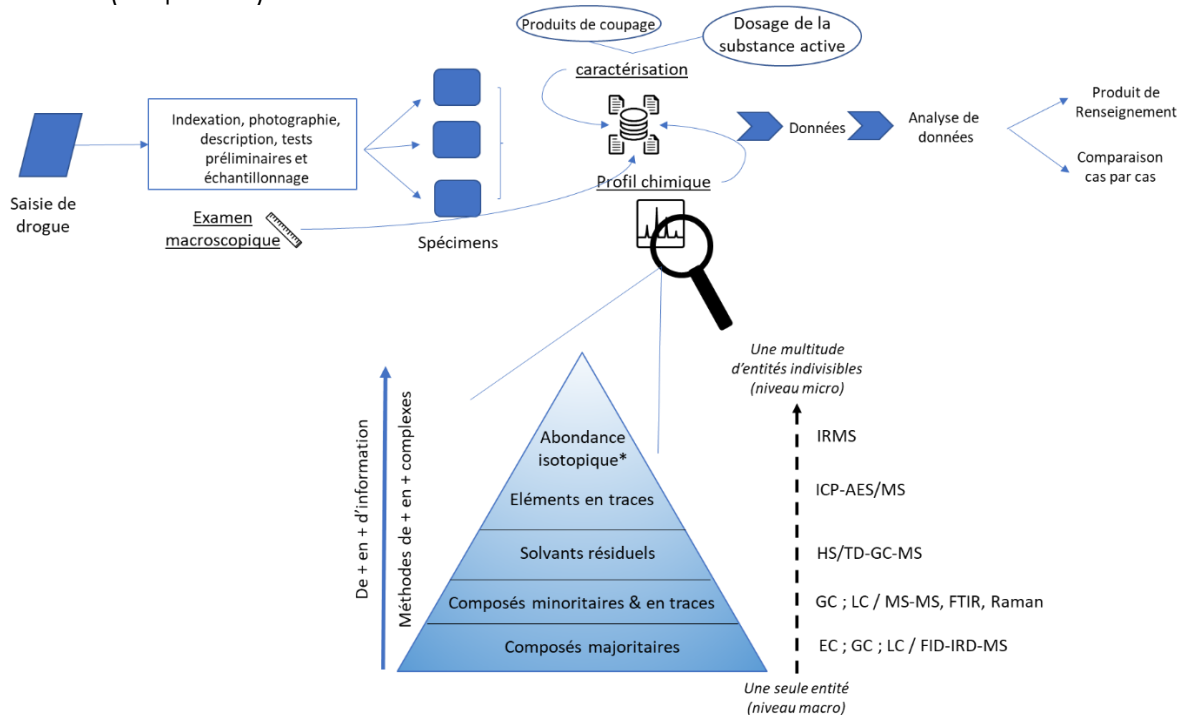


Figure 25 : Illustration du processus général de profilage. IRMS = Spectrométrie de masse à rapport isotopique ; ICP-AES/MS = plasma à couplage inductif (ICP) couplé à un Spectromètres à Emission Atomique (AES) ou à un spectromètre de masse (MS) ; HS/TD-GC-MS = Systèmes espace de tête statique (HS) et thermodésorption (TD) couplés chromatographie gazeuse (GC) et spectrométrie de masse (MS) ; GC ; LC / MS-MS, FTIR = chromatographie gazeuse (GC) ou liquide (LC) couplée à un spectromètre de masse en tandem. Spectroscopie infrarouge à transformée de Fourier ; EC ; GC ; LC / FID-IRD-MS = électrophorèse capillaire, GC ou LC couplée à une détection à ionisation de flamme, détecteur infrarouge, à un spectromètre de masse.

*L'analyse isotopique, qui est une méthode complexe, ne cible pas le niveau micro (échelle de droite), car seuls les rapports isotopiques des principaux composants sont mesurés et ne donne pas beaucoup d'informations (échelle de gauche) ; elle améliore la capacité de discrimination des analyses de géolocalisations existantes.

La sélection des caractéristiques à prendre en compte pour un profil est un long cheminement. Seules celles qui sont stables dans le temps (dont l'abondance ne varie pas en fonction de la durée de stockage ou les coupages successifs par exemple) et qui sont pertinentes (si un composé est systématiquement présent dans les mêmes proportions, il ne présente pas d'intérêt, car il est peu discriminant) seront choisies pour constituer l'empreinte chimique. Elles peuvent être qualitatives (présence ou absence d'un composé) et/ou quantitatives (abondance relative d'un composé).

On peut donc imaginer un profil chimique comme étant un ensemble de variables (par exemple des aires de pic chromatographique comme sur la figure 26 ci-dessous) et la comparaison entre 2 profils résulte de la superposition de ces 2 ensembles pour établir un « match » (même principe que la comparaison de profils génétiques).

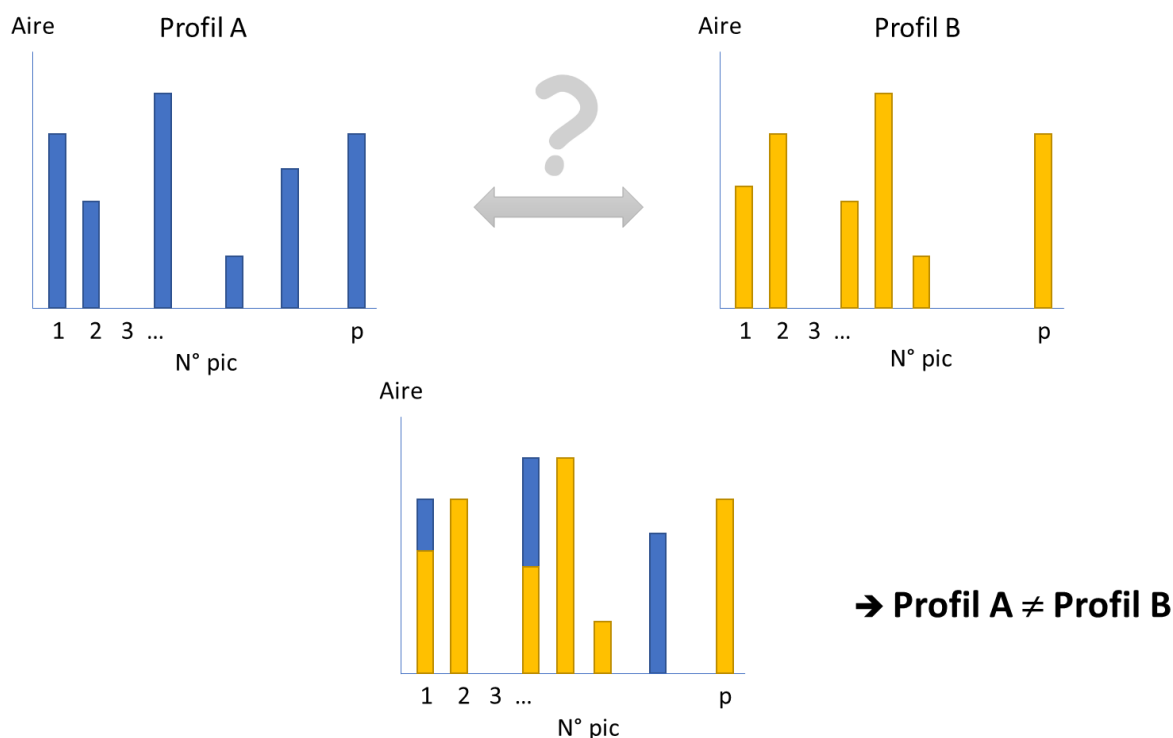


Figure 26 : Principe de comparaison de 2 profils chimiques

L'établissement d'un profil chimique pour réaliser des comparaisons est assorti d'un traitement statistique qui permet de déterminer de façon fiable et avec un pourcentage de risque (incertitude) le degré de similitude de 2 profils. C'est à dire, à partir d'un certain seuil de similitude, 2 échantillons seront déclarés liés (issus d'un même lot de fabrication par exemple) ; en dessous d'un autre seuil, les échantillons seront déclarés non liés.

Le terme « signature chimique », est généralement ce que l'on entend par « profilage de drogues »⁴⁴.

2.6.2. Genèse du profilage chimique des drogues

Pendant des décennies, le trafic de cocaïne, d'héroïne a donné lieu à des activités criminelles à grande échelle dans de nombreux pays. Dans les années 1960 et 1970, le trafic d'héroïne et la dépendance à cette substance ont augmenté à un point tel qu'ils sont devenus un problème mondial grave, suscitant l'inquiétude de nombreux

gouvernements. En 1977, le Programme des Nations Unies pour le contrôle international des drogues a convoqué une réunion spéciale d'experts pour discuter de la caractérisation chimique ou « profilage » de l'héroïne. En 1982 un programme de recherche sur le profilage chimique de l'héroïne dont l'objectif était d'utiliser la chimie pour la géolocalisation a été élaboré. En 1992, il a été décidé d'étendre la portée du programme à la cocaïne et aux amphétamines. L'accent a été porté sur le profilage des impuretés par chromatographie en phase gazeuse couplée à un détecteur de masse (GC-MS), c'est-à-dire sur l'identification des impuretés alcaloïdes coextraites du pavot à opium et de la feuille de coca, ainsi que des sous-produits de la transformation en chlorhydrate d'héroïne et chlorhydrate de cocaïne. En 2005, un manuel contenant des méthodes recommandées pour le profilage des impuretés de l'héroïne et de la cocaïne a été publié par l'ONUDC en collaboration avec de nombreux services de police scientifique gouvernementaux de par le monde ⁴⁵.

La recherche sur les techniques de profilage chimique a été stimulée par les attaques terroristes du 11 septembre 2001. De nombreuses agences de renseignement pensaient que les activités terroristes étaient, en partie, financées par le trafic de drogue ⁴⁴.

Le profilage a été limité à 5 drogues illicites principalement : l'héroïne, la cocaïne, l'amphétamine, la méthamphétamine et la MDMA, principalement parce qu'elles sont produites en grandes quantités et associées à des dommages sociaux importants (accidents et suicides, criminalité et violence, dépendance ⁴⁶).

De nombreux laboratoires de police scientifique dans le monde, majoritairement des laboratoires publics à un niveau ou à un autre, disposent d'une capacité d'analyse permettant de réaliser les 2 premières étapes d'identification et de quantification. Cependant, très peu disposent de programmes de profilage de drogues complets et spécialisés.

Si le profilage chimique des drogues illicites a d'abord été établi pour fournir des renseignements stratégiques, par exemple la géolocalisation, ces mêmes informations chimiques ont également été utilisées pour établir des liens entre les échantillons. En effet, la détermination de l'origine géographique nécessite d'avoir une base de données de spécimens provenant de sources géographiques connues, c'est-à-dire saisis directement dans la zone de production et dans des laboratoires clandestins ce qui est très difficile à obtenir.

Si la recherche dans le domaine du profilage a été longtemps dominée par les États-Unis, au niveau européen, des études ont été menées notamment depuis 1993 en Suisse ⁴⁷ et en 1999 en France ⁴⁸.

À partir de 1999, plusieurs projets ont été financés par la Commission Européenne. Ils seront détaillés ultérieurement.

2.6.3. Principe de la comparaison de profils chimiques

Un profil chimique peut être considéré comme un vecteur avec autant de variables qu'il y a de composés cibles. Les vecteurs sont ensuite stockés dans un tableau (ou dans une base de données) pour être ensuite comparés entre eux à l'aide de méthodes d'analyses multivariées. Les méthodes utilisées dépendent de l'information souhaitée.

On peut utiliser soit des méthodes de réduction de dimension comme l'analyse en composantes principales ou des mesures de distances (figures 27a et 27b).

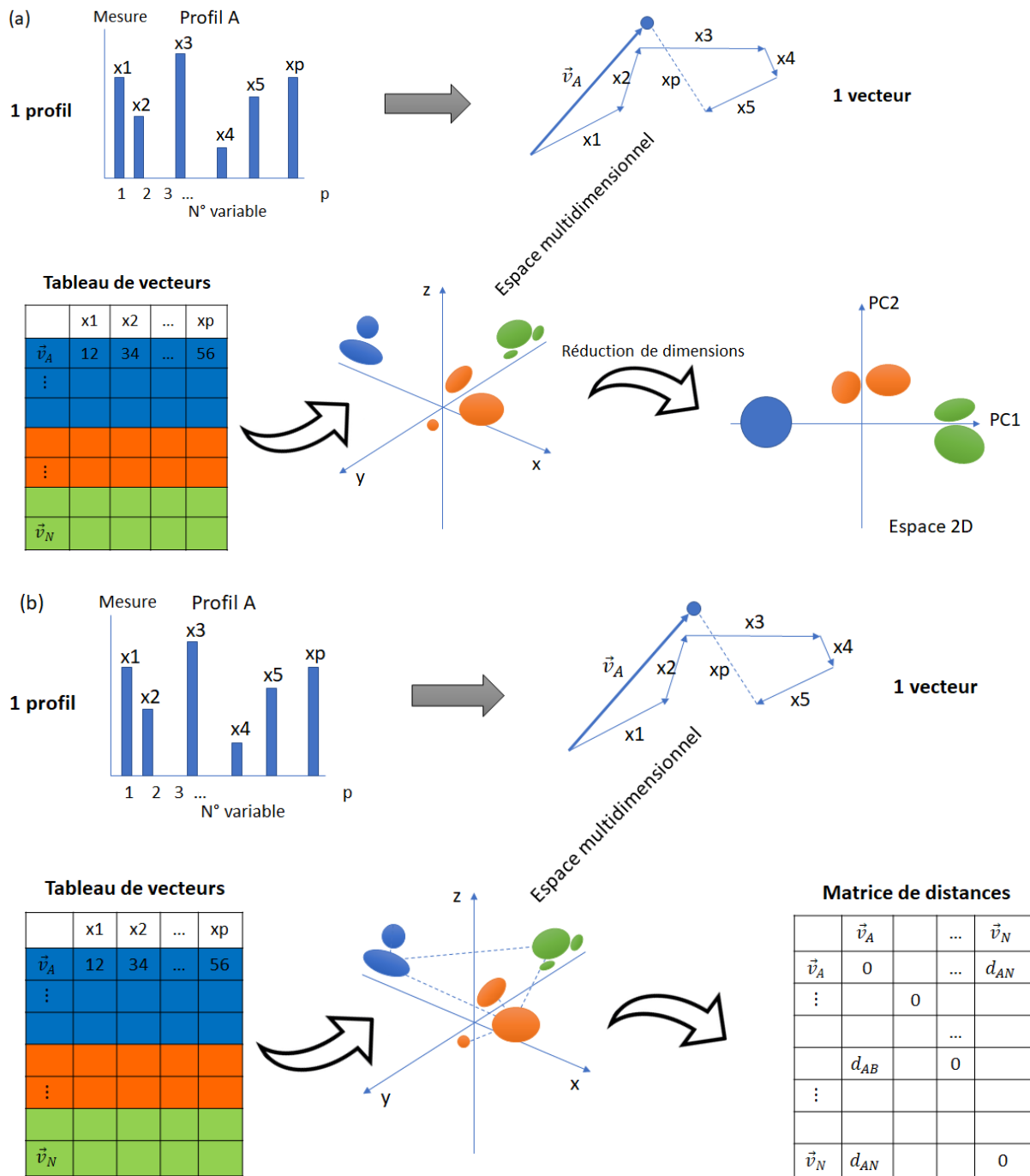


Figure 27 : Méthodes d'analyses multivariées : (a) Réduction de dimensions. (b) mesure de distances.

En analyse des données, la mesure de similarité est un moyen de mesurer la façon dont des individus (ou observations ou objets) sont liés ou proches les uns des autres. La mesure de dissimilarité est de dire à quel point les objets sont distincts. Ces termes sont souvent utilisés dans les méthodes de clustering lorsque des objets similaires sont regroupés en un seul cluster (ou groupe). Il est également utilisé dans la classification (par exemple en utilisant les K plus proches voisins ou K-Nearest Neighbors (kNN)), où les objets sont étiquetés en fonction de la similarité de leurs caractéristiques. Un autre exemple est la détection d'anomalie ou de points aberrants c'est-à-dire des points très différents par rapport à l'ensemble des objets.

Une mesure de similarité est une mesure numérique de la similitude de 2 objets. Elle est souvent comprise entre 0 (aucune similitude) et 1 (similitude totale).

Une mesure de dissimilarité est une mesure numérique de la différence entre 2 objets. Elle est comprise entre 0 (les objets sont semblables) et 1 (les objets sont différents).

Une proximité fait référence à une similarité ou une dissimilarité.

Une distance est une mesure de dissimilarité et est censée être une métrique si et seulement si elle satisfait les 4 conditions suivantes ^{49,50} :

1. non-négativité : $d(A, B) \geq 0$, pour 2 observations distinctes quelconques A et B
2. symétrie : $d(A, B) = d(B, A)$ pour tout A et B
3. inégalité triangulaire : $d(A, B) \leq d(A, C) + d(C, B)$ pour tout A, B, C
4. $d(A, B) = 0$ seulement si $A = B$

La technique utilisée pour mesurer les distances dépend de la situation dans laquelle on travaille. Par exemple, dans certains cas, la distance euclidienne peut être optimale et utile pour calculer les distances. D'autres applications nécessitent une approche plus sophistiquée pour calculer les distances entre les points ou les observations comme la mesure angulaire à l'aide du cosinus (ou du sinus).

La figure 28 illustre la différence entre 4 types de distance, Euclidienne, de Manhattan, de Chebychev et de Minkowski :

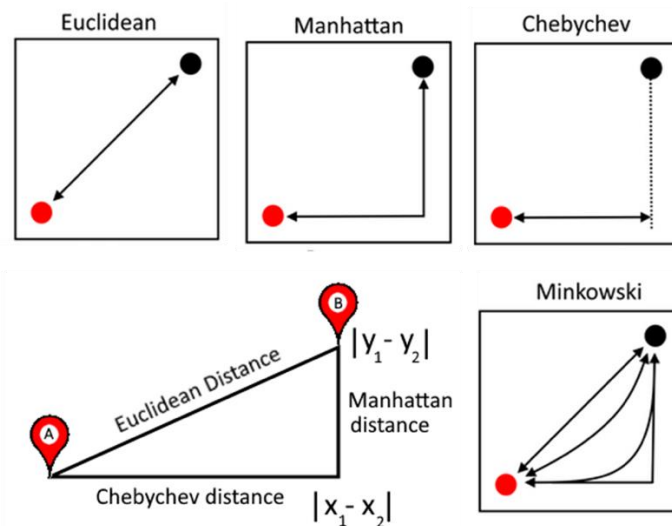


Figure 28 : Différences entre les distances Euclidienne, de Manhattan, de Chebychev et de Minkowski (d'après <https://towardsdatascience.com/17-types-of-similarity-and-dissimilarity-measures-used-in-data-science-3eb914d2681>).

Le tableau 4 ci-après répertorie quelques métriques ^{51,52} :

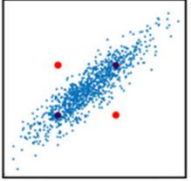
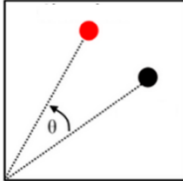
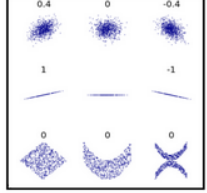
Métrique	Equation	Avantages	Inconvénients
Distance de Minkowski	$d_{Min} = \left(\sum_{i=1}^n X_i - Y_i ^n \right)^{\frac{1}{n}}$	Généralisation à la fois de la distance euclidienne et de la distance de Manhattan	Sensible aux valeurs aberrantes
Distance de Manhattan	$d_{Man} = \sum_{i=1}^n X_i - Y_i $ égale à la distance de Minkowski pour n = 1	Est commun et comme les autres distances pilotées par Minkowski. Fonctionne bien avec des ensembles de données avec des clusters compacts ou isolés	Sensible aux valeurs aberrantes
Distance euclidienne	$d_{euc} = \sqrt{\sum_{i=1}^n (X_i - Y_i)^2}$ égale à la distance de Minkowski pour n = 2	Très commune, facile à calculer et fonctionne bien avec des ensembles de données avec des groupes compacts ou isolés	Sensible aux valeurs aberrantes
Distance de Chebichev	$d_{Man} = \max_{i=1}^n (X_i - Y_i)$	Moins courante. Distance entre 2 points donnée par la différence maximale entre leurs coordonnées sur une dimension	Sensible aux valeurs aberrantes
Distance de Mahalanobis	$d_{Mah} = \sqrt{(X - Y)S^{-1}(X - Y)^T}$ utilisée dans les clusters ellipsoïdes, S est la matrice de covariance	L'utilisation la plus courante de cette distance consiste à trouver des valeurs aberrantes multivariées, ce qui indique des combinaisons inhabituelles de 2 variables ou plus	<p>Mahalanobis</p>  <p>Peut être coûteuse en termes de calcul</p>
Distance de Canberra	$d_{can} = \sum_{i=1}^n \frac{ X_i - Y_i }{(X_i + Y_i)}$	Résultats précis pour les ensembles de données à haute dimension à l'aide de l'algorithme K-médoides	-
Fonction cosinus (similarité)	$Cos(x, y) = \frac{\sum_{i=1}^n X_i Y_i}{\ X\ _2 \ Y\ _2}$	Indépendant de la longueur du vecteur et invariant à la rotation	<p>Cosine</p>  <p>N'est pas invariant à la transformation linéaire</p>
Coefficient de Pearson	$\rho(x, y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$	Indépendant de la longueur du vecteur et invariant à la rotation	<p>Pearson</p>  <p>N'est pas invariant à la transformation linéaire</p>

Tableau 4 : mesures de distances et de similarités.

On verra par la suite dans quel cas ces méthodes sont utilisées.

En général, les données brutes sont rarement utilisables telles quelles en raison de facteurs inhérents à la préparation de l'échantillon et à son analyse. Il est donc nécessaire d'en tenir compte lorsqu'on souhaite faire des comparaisons.

En chromatographie, par exemple, les différences d'abondance globale entre les chromatogrammes mesurés sur le même instrument peuvent résulter bien sûr d'une variation dans la préparation de l'échantillon, mais également de l'injection de l'échantillon, des conditions chromatographiques et de la réponse de l'instrument.

L'autre phénomène à prendre en compte est l'effet de dilution : l'aire d'un pic chromatographique est directement liée à la concentration d'un composé dans le mélange. Les échantillons provenant de saisies de rue ont en général des concentrations en principe actif plus faibles que ceux provenant d'une saisie réalisée plus haut dans la chaîne de distribution (au niveau du dealer par exemple). Par conséquent si on souhaite comparer ces échantillons entre eux, il faudra tenir compte de cet effet de dilution sur toutes les aires des composés cibles (cf. figure 29).

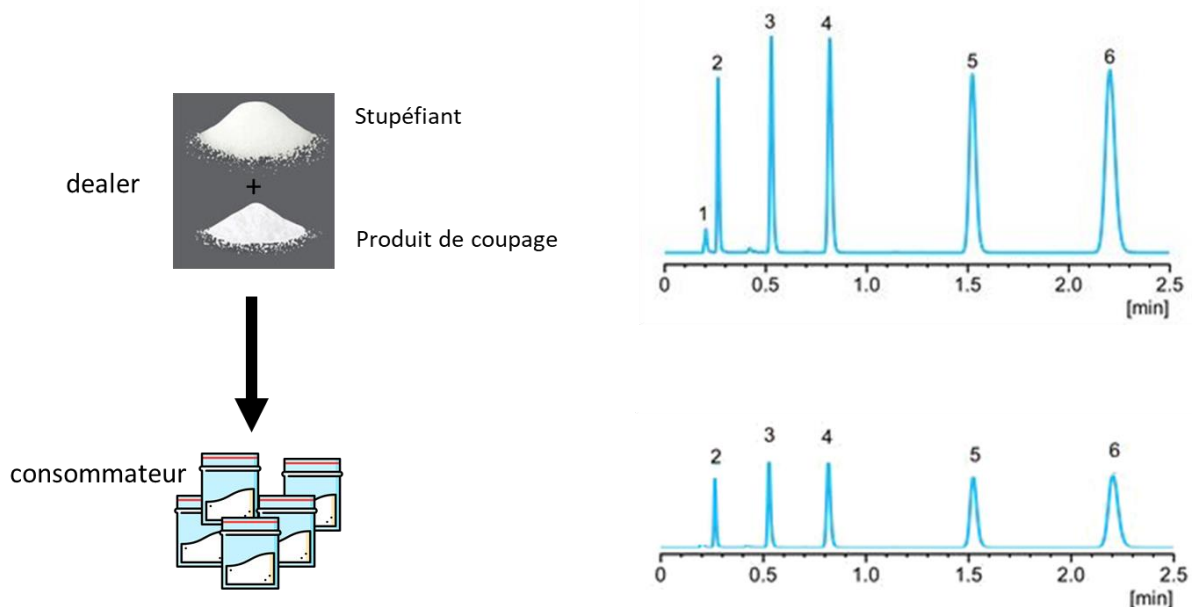


Figure 29 : Effet de la dilution des drogues sur la forme du chromatogramme.

Afin de corriger ces problèmes, des procédures de prétraitement des données initiales sont couramment appliquées. Il faut toutefois garder à l'esprit que le prétraitement modifie les données, ce qui influence positivement ou négativement les résultats. Pour que l'analyse globale des données soit réussie, il est essentiel de "se préparer" en appliquant le ou les outils de prétraitement appropriés. La sélection du prétraitement optimal peut nécessiter une certaine itération entre l'analyse de données primaire et l'étape de prétraitement. Bien que cette approche empirique soit une pratique courante, il est préférable que l'outil de prétraitement soit choisi en raison d'une caractéristique connue des données ⁵³.

Dans l'exemple figure 29, une normalisation par rapport à l'aire du principe actif (héroïne, amphétamine, etc.) permettra de prendre en compte la différence de concentration dans les chromatogrammes, mais des pics d'impuretés peuvent avoir

disparu sous l'effet de la dilution et ils ne pourront pas être « récupérés », c'est pourquoi il est nécessaire d'avoir une connaissance approfondie de la fabrication des produits analysés.

2.6.3. Évaluation des mesures de dissimilarités

Comment savoir si une métrique est meilleure qu'une autre pour différencier des échantillons qui sont supposés être différents et pour rapprocher ceux qui sont censés être similaires ?

Il est donc nécessaire d'évaluer leur capacité de discrimination : on s'attendra à obtenir des distances proches de 0 si les échantillons comparés entre eux sont issus d'un même lot de fabrication et des distances très différentes de 0 s'ils proviennent de lots différents. Mais comment savoir à quel moment on peut faire la différence ?

Dans un premier temps, on calcule les distances de tous les échantillons liés et celles de tous les non liés pour construire 2 ensembles : un ensemble de distances appelées intralots et un ensemble de distances appelées interlots. On représente ensuite les distributions de ces 2 ensembles de distances. On obtient le type de graphique illustré dans la figure suivante (figure 30) :

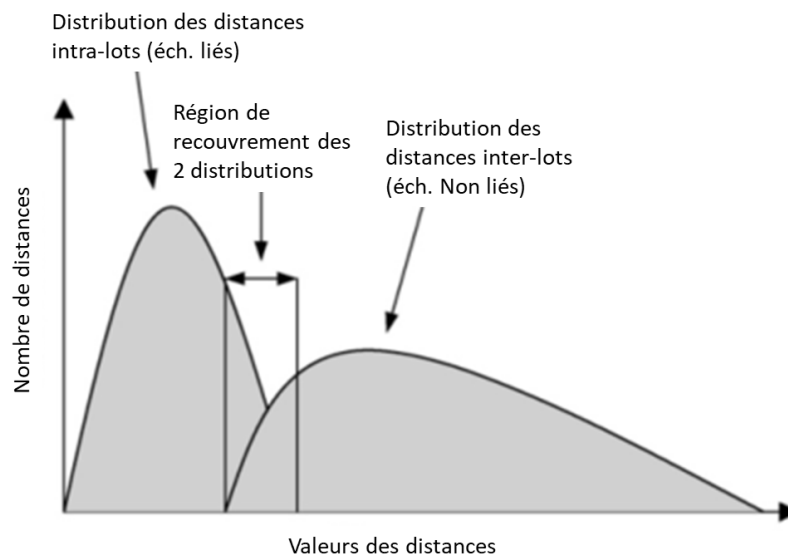


Figure 30 : Visualisation des distributions des distances des 2 ensembles d'échantillons : liés (distances intralots) et non liés (distances interlots). La zone entre les 2 est appelée zone ou région de recouvrement.

Plus les 2 distributions sont éloignées, plus petite est la région de recouvrement et par conséquent meilleure est la discrimination.

Si on considère, dans un premier temps, que les distributions associées à des populations peuvent être approximées par des lois normales, on peut utiliser les paramètres de la moyenne et de l'écart-type pour calculer un rapport appelé quotient ou pouvoir de discrimination (DP pour discriminating power) (équation 3) ⁵⁴ qui vaut :

$$DP = \frac{\mu_{non\ liés} - k\sigma_{non\ liés}}{\mu_{liés} + k\sigma_{liés}}$$

Équation 3

Avec k = 1, 2 ou 3 suivant le niveau de confiance considéré

Ce quotient est une mesure de la capacité des différents prétraitements et métriques à séparer les 2 distributions. Plus il est élevé, meilleure est la séparation. L'inférence statistique et ici paramétrique (distributions supposées normales) est à la base de ce calcul.

Dans un deuxième temps, une fois la méthode de prétraitement et la métrique sélectionnées, il est alors possible d'établir un seuil pour lequel on calculera les pourcentages de faux positifs et de faux négatifs, par analogie avec les tests de dépistage de maladie (cf. figure 31).

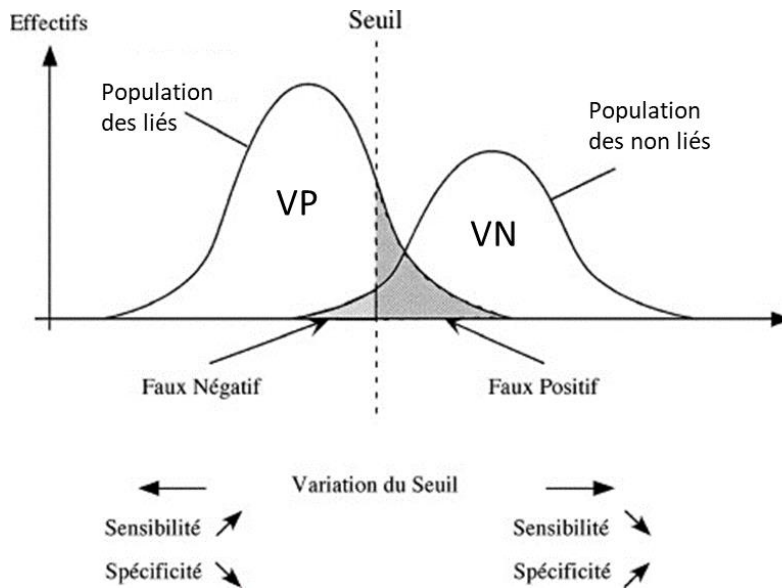


Figure 31 : Représentation du calcul des faux positifs et des faux négatifs. VN = vrais objets non liés ; VP = vrais objets liés.

Une méthode de profilage est d'autant plus performante que son pourcentage de faux positifs (FP) sera faible. En effet, un faux positif correspond à déclarer qu'un spécimen d'une saisie est lié à un spécimen d'une autre saisie alors qu'il n'en est rien. Cette déclaration peut conduire les enquêteurs sur de mauvaises pistes, il est donc préférable de limiter le plus possible ce type d'erreur. Il est moins préjudiciable d'avoir un taux de faux négatifs (FN) élevé : on déclare qu'un spécimen d'une saisie est non lié à un spécimen d'une autre saisie alors qu'ils le sont, cela revient à ne pas fournir d'information.

Pour illustrer le calcul des indices de sensibilité et de spécificité, on présente les résultats sous la forme d'un tableau (tableau 5) :

Méthode de profilage	Spécimens « liés »	Spécimens « non liés »	Total
On déclare lié	VP	FP	VP+FP
On déclare non lié	FN	VN	FN+VN
Total	VP+FN	FP+VN	VP+VN+FP+VN

Tableau 5 : tableau pour le calcul des indices de spécificité et de sensibilité

Calcul de la sensibilité : $Se = VP / (VP + FN)$ (vrais positifs divisés pas le nombre de spécimens effectivement "liés")

Calcul de la spécificité : $Sp = VN/(FP+VN)$ (vrais négatifs divisés par le nombre des spécimens effectivement "non liés").

Plus le seuil est élevé, plus la sensibilité diminue, on aura tendance à ne pas faire de liens entre affaires et plus le seuil est faible, plus la sensibilité augmente, on fera plus de liens entre affaires, mais beaucoup seront erronés.

Le choix du seuil est donc un critère très important. On peut aussi s'aider de courbes ROC (Receiver Operating Characteristic) qui représentent la sensibilité en fonction de 1 – spécificité pour toutes les valeurs seuils possibles afin d'évaluer celui qui est optimal.

Apport de l'approche bayésienne

Quelles que soient les données (drogues illicites ou autre type de données), la classification des échantillons et l'établissement d'un lien sont basés sur un seuil et les résultats sont susceptibles d'être interprétés de manière dichotomique.

Cette procédure peut être critiquée sur 2 points. Premièrement, dans l'étape de comparaison, le seuil (S) constitue une étape qualitative qui peut donner des résultats très différents pour 2 paires d'observations différentes. Une paire d'observations peut fournir une valeur de S qui est juste en dessous du seuil, alors que l'autre paire peut fournir une valeur de S juste au-dessus du seuil. La première paire passera à l'étape de différence significative, la seconde non. Pourtant, les 2 paires peuvent avoir des mesures qui sont proches. La différence dans les conséquences est plus grande que la différence dans les mérites des mesures. Une telle approche est appelée effet « chute de la falaise ou fall-off-the-cliff »). Ce changement soudain de décision dû au franchissement d'une ligne particulière est comparé à la chute d'une falaise, un moment vous êtes en sécurité, le suivant mort. En fait, plutôt qu'une falaise, on a simplement une pente raide ⁵⁵. Toutes choses étant égales par ailleurs, plus les échantillons sont similaires, plus la preuve qu'ils ont une origine commune est forte et moins les échantillons sont similaires, plus la preuve qu'ils proviennent de sources différentes est forte ⁵⁶.

La deuxième critique est que le résultat est difficile à interpréter. En raison de l'effet de l'étape de comparaison, le résultat n'est pas simplement la probabilité de la preuve, en supposant que les 2 observations proviennent de sources différentes.

La première étape écarte certains ensembles de mesures qui peuvent provenir de la même source et n'écarte pas nécessairement d'autres ensembles de mesures qui proviennent de sources différentes. La deuxième étape calcule une probabilité, non pas de la preuve, mais de la partie de la preuve pour laquelle S n'était pas supérieur à la valeur seuil, en supposant que les 2 ensembles provenaient de sources différentes. Il est nécessaire de comparer cette probabilité avec la probabilité du même résultat, en supposant que les 2 ensembles proviennent de la même source. La détermination de la probabilité au stade de la source implique également qu'une faible probabilité pour la preuve, en supposant que les deux ensembles proviennent de sources différentes, signifie qu'il existe une forte probabilité que les deux ensembles proviennent de la même source. Cette implication n'est pas fondée ⁵⁵.

Une autre **approche basée sur le rapport de vraisemblance** (likelihood ratio ou LR) du théorème de Bayes, **fournit une mesure de la valeur de la preuve (ici le lien chimique)** qui diminue à mesure que la distance entre les 2 ensembles de mesures augmente, sous réserve, de la rareté ou non des mesures.

Lorsque l'évaluation de la preuve est basée sur des données analytiques obtenues à partir d'une analyse physico-chimique, une attention particulière aux considérations suivantes est requise ⁵⁷ :

1. les sources possibles d'incertitude (sources d'erreur), qui doivent au moins inclure les variations dans les mesures des caractéristiques au sein des objets retrouvés et/ou des objets de contrôle, et les variations dans les mesures des caractéristiques entre divers objets dans la population pertinente (par ex. la population des saisies d'un stupéfiant en particulier),
2. des informations sur la rareté des caractéristiques physico-chimiques déterminées (par exemple, la composition élémentaire et/ou chimique des échantillons comparés) pour les échantillons saisis et/ou de contrôle dans la population pertinente,
3. le niveau d'association (corrélation) entre les différentes caractéristiques lorsque plus d'une caractéristique a été mesurée,
4. dans le cas du problème de comparaison, la similarité du matériau retrouvé et de l'échantillon de contrôle.

Une première façon d'interpréter les résultats des comparaisons de saisies d'héroïne utilisant une approche continue a été proposée en 2003. Une seconde façon a été proposée pour la comparaison des saisies de comprimés de MDMA en 2009 (cf. § Approche bayésienne du profilage des impuretés).

L'évaluation du lien peut être abordée de manière similaire à celle des fragments de verre. Dans les cas impliquant des fragments de verre, si des fragments sont retrouvés sur les vêtements d'un suspect, le scientifique cherche à savoir si ceux-ci pourraient provenir d'un échantillon témoin (une fenêtre brisée). Dans une telle situation, la source potentielle est représentée par le verre de la fenêtre ⁵⁸.

Dans le cas des échantillons de drogues illicites, la source est généralement inconnue (le laboratoire où la drogue est fabriquée est inconnu). Par conséquent, la connaissance de la variabilité chimique du lot de drogue fabriqué et mis sur le marché est incomplète (incertaine).

Pour évaluer la valeur du lien entre 2 saisies λ et γ , on a utilisé le rapport de vraisemblance $LR = P(E|H_1)/P(E|H_2)$, pour lequel les hypothèses H_1 correspondent à : « La saisie λ provient de la même source que la saisie γ », et l'alternative H_2 correspond à : « La saisie λ provient d'une source différente de la saisie γ ». La preuve E est la corrélation des résultats d'analyse des échantillons des deux saisies.

Profilage des impuretés de l'amphétamine

Les travaux sur ce sujet ont été réalisés dans le cadre d'un projet paneuropéen (SMT-CT98- 2277) d'une durée de 3 ans, financé par le programme de recherche de la DG XII (Normes, mesures et essais) de la Commission européenne et l'Office fédéral de l'éducation et de la science de la Suisse ⁵⁹.

Mon postdoctorat à l'université de Lausanne entre 2000 et 2002 était financé par ce projet. Il était divisé en 6 parties :

- I. Synthèse des standards et compilation des données analytiques ⁶⁰,
- II. Stabilité des impuretés dans les solvants organiques ⁶¹,
- III. Développement de la méthode en chromatographie gazeuse (GC) ⁶²,
- IV. Optimisation de la préparation d'échantillon ⁶³,

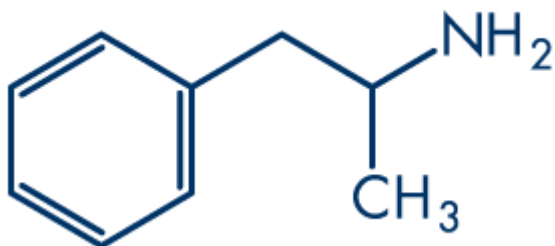
V. Détermination de la variabilité de la méthode optimisée ⁶⁴,

VI. Évaluation des méthodes de comparaison de l'amphétamine ⁶⁵.

Carte d'identité de l'amphétamine

L'amphétamine (N-**Alpha-MéthylPHénETHYLAMINE**) est une substance de synthèse, chef de file d'une liste comprenant quelques centaines de molécules chimiquement proches les unes des autres.

Structure moléculaire :



Sa formule chimique est : C₉H₁₃N de masse moléculaire 135.2 g/mol.

L'atome de carbone α asymétrique donne lieu à deux énantiomères définis comme les stéréo-isomères R et S^m.

Les amphétamines ou « speed » font partie de la famille des phénéthylamines.

L'amphétamine base est une huile volatile incolore insoluble dans l'eau. Le sel le plus courant est le sulfate qui se présente sous forme d'une poudre blanche ou blanc cassé soluble dans l'eau. Les produits illicites sont généralement constitués de poudres. L'amphétamine peut être ingérée, sniffée et, moins fréquemment, injectée. Contrairement au sel chlorhydrate de méthamphétamine, le sulfate d'amphétamine n'est pas assez volatil pour être fumé. En cas d'ingestion, la dose peut varier entre plusieurs dizaines et plusieurs centaines de milligrammes en fonction de la pureté. La dose létale minimale chez les adultes non dépendants est estimée à 200 mgⁿ.

^m https://www.emcdda.europa.eu/publications/drug-profiles/amphetamine_en

ⁿ https://www.emcdda.europa.eu/publications/drug-profiles/amphetamine_fr consulté le 18/05/2022

Petite histoire des amphétamines

La première purification chimique d'une amphétamine, l'éphédrine à partir du mahuang ou Ma Huang, plante utilisée depuis plusieurs millénaires en Chine est attribuée à Nagai Nagayoshi, un pharmacien japonais et date de 1885 (Lee 2011). Au même moment, en Allemagne, un chimiste, d'origine roumaine, Lazăr Edeleanu, réalise la première synthèse d'amphétamines. Mais sa découverte est peu remarquée. La synthèse de l'ecstasy (ou MDMA, pour 3,4-méthylènedioxy-N-méthylamphétamine) par les laboratoires Merck remonte quant à elle à 1912, et la première synthèse d'une cathinone (la β -kétamphétamine) à 1929 (Freudenmann, Oxler, et Bernschneider-Reif 2006).

Dans la première moitié du XXe siècle, les effets de ces molécules de synthèse ont captivé les chimistes et suscité l'intérêt du monde médical. Leurs activités sympathomimétique (stimulant), anorexigène (coupe-faim) et psychostimulante (dopante) ont conduit à la commercialisation des premiers médicaments à base d'amphétamine : un bronchodilatateur avec la Benzédrine® en 1934 aux États-Unis (un sulfate d'amphétamine), un coupe-faim avec l'Obetrol® dès les années 1950 également aux États-Unis (une combinaison de sels d'amphétamine).

C'est en partie grâce aux amphétamines que les athlètes américains écrasent ceux de l'Allemagne nazie lors des jeux de Berlin de 1936. Cette humiliation provoque l'accélération des recherches par les chimistes du Reich qui synthétisent, en 1937, une amphétamine encore plus efficace. Il s'agit de la méthamphétamine, qui sera commercialisée sous le nom de Pervitine. Encore appelée « pilule de Göring », cette préparation est administrée massivement sous forme de comprimés aux soldats allemands tout au long de la Seconde Guerre mondiale. Des historiens ont estimé que sans elle la stratégie de guerre éclair (« blitzkrieg ») de l'Allemagne en Pologne ou en France n'aurait pas été possible. Les aviateurs américains et britanniques, eux aussi, ont utilisé des amphétamines pour améliorer leur vigilance lors de leurs sorties.

Ces substances ont continué à être utilisées par les armées du monde entier pour doper leurs aviateurs – guerres du Golfe, Afghanistan... Plus récemment, les médias ont désigné le captagon, substance apparentée aux amphétamines, comme « la drogue des jihadistes » de Daesh. Dans un autre registre, les amphétamines ont aussi accompagné l'aventure de la conquête spatiale, aussi bien du côté russe que du côté américain. L'équipage de la mission Apollo 13, par exemple, en a consommé pour tenir le coup lors de son retour vers la Terre après l'accident survenu à leur vaisseau.

En dehors de la sphère militaire, la consommation d'amphétamines s'est banalisée dans les années d'après-guerre. Elles étaient alors disponibles sans ordonnance. Dans ces années de reconstruction, il était courant d'en prendre avant une grosse journée de travail ou avant un examen aussi bien en Europe qu'aux États-Unis ou au Japon. Les cadres japonais en ont particulièrement abusé dans les années 50 .

Leur toxicité ainsi que la forte dépendance psychique qu'elles entraînent expliquent que la plupart des amphétamines soient classées en France comme stupéfiants depuis 1967.

La principale voie de synthèse fait appel au procédé de **Leuckart**. Celui-ci utilise la 1-phényl-2-propanone (**P2P** ou **BMK** ou phénylacétone) ainsi que des réactifs tels que l'acide formique, le formate ou le formamide d'ammonium pour produire un mélange racémique des énantiomères R et S. La voie de l'amination réductrice utilise également la BMK comme précurseur (cf. figure 32)⁶⁶.

Un procédé nettement moins courant, mais stéréosélectif est basé sur la réduction des diastéréoisomères de la **noréphédrine** ou de la **norpseudoéphédrine**⁶⁷.

Ces précurseurs (1-phényl-2-propanone, noréphédrine ou norpseudoéphédrine) sont répertoriés au tableau I de la Convention de 1988 des Nations unies contre le trafic illicite des stupéfiants et des substances psychotropes.

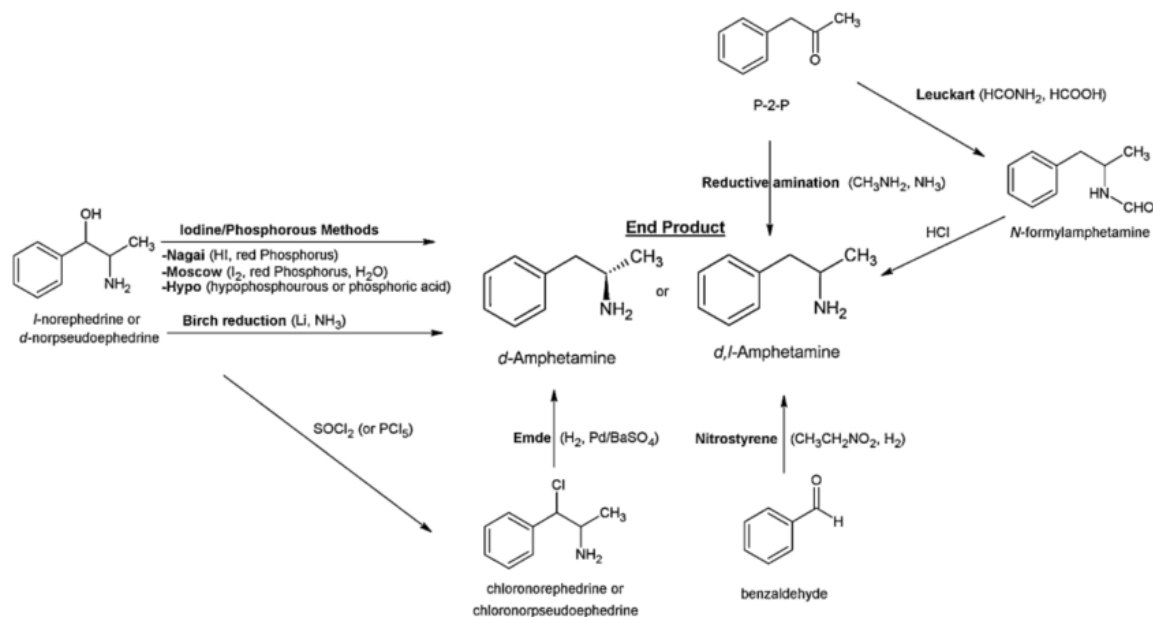


Fig. 3. Common synthetic routes of AP manufacture.

Figure 32 : Voies de synthèse communes de l'amphétamine (d'après Stojanovska et al. 66).

La dernière partie du projet (partie VI) traitait de l'évaluation des méthodes statistiques de comparaison de différents types d'amphétamines que ce soit au niveau des synthèses ou des lots de fabrication (différentes recettes).

Le profilage de l'amphétamine a été réalisé principalement par chromatographie en phase gazeuse (GC) couplée à la spectrométrie de masse (MS).

Quarante-quatre échantillons issus de lots de fabrication bien définis ont été utilisés dans l'étude, et ils ont été choisis pour être représentatifs d'un total de 768 disponibles et parfaitement contrôlés. Un échantillon correspond à un prélèvement d'une certaine quantité de poudre dans un lot. Pour chaque échantillon, 3 prélèvements ont été réalisés (correspondant à des répétitions).

Les échantillons sélectionnés ont été synthétisés suivant 3 voies :

1. réaction de Leuckart (25 échantillons de 25 lots différents),
2. amination réductrice de la BMK (12 échantillons de 12 lots différents),
3. voie du nitrostyrène (7 échantillons de 7 lots différents).

Dans une première phase de sélection, 33 sous-produits (des composés cibles) ont été choisis en raison des principaux critères suivants :

1. rencontrés dans les amphétamines de rue,
2. stables en solution,
3. surfaces de pic reproductibles.

Les profils chimiques des 44 amphétamines, représentés comme des vecteurs à 33 variables maximum (toutes les variables n'étaient pas présentes) ont été stockés dans un tableau.

L'effet de la réduction de ce nombre de composés cibles sur la séparation de différents échantillons a été étudié en effectuant une analyse discriminante partielle des moindres carrés (PLS-DA) et des calculs de distance.

Sept méthodes de distance et de similarité ont été testées et évaluées quant à leur capacité à distinguer les échantillons liés (distances entre échantillons du même lot) des échantillons non liés (distances entre échantillons de lots différents) : la distance de Manhattan, la distance euclidienne, la distance de Canberra, l'indice de similarité (équation 4), la corrélation de Pearson, le sinus au carré (équation 5) et la méthode du quotient ⁵⁴.

Pour permettre la comparaison des résultats, les distances obtenues avec chaque méthode ont été organisées pour fournir des valeurs entre 0 et 100, où 0 indique la comparaison de profils identiques.

$$SI = \frac{100 \times 50}{n} \sum_{i=1}^n \left(\left(\frac{X_i}{Y_i} - 0.25 \right)^6 - 1 + 50 \right)^{-1}$$

Si $X_i > Y_i$ alors $\frac{X_i}{Y_i}$ et si $X_i < Y_i$ alors $\frac{Y_i}{X_i}$

Équation 4

$$(\sin \theta)^2 = 1 - \left(\frac{\sum_{i=1}^n (X_i Y_i)}{\sqrt{\sum_{i=1}^n X_i^2 \sum_{i=1}^n Y_i^2}} \right)^2$$

Équation 5

Des prétraitements ont été utilisés avant de comparer les échantillons entre eux : des combinaisons de pondération, de normalisation et d'utilisation du logarithme et de la racine quatrième. La pondération a été effectuée en divisant la surface du pic de chaque composé cible par son écart-type (S.D.). L'écart-type de chaque composé a été calculé sur l'ensemble de données des 768 échantillons. La normalisation a été réalisée en divisant la surface du pic de chaque composé cible dans le chromatogramme par la somme des surfaces des pics de tous les composés cibles dans le même chromatogramme.

Deux des méthodes de distance, à savoir la distance de Canberra et l'indice de similarité, ne pouvaient pas gérer la présence de zéros dans le tableau de données, c'est-à-dire lorsque les composés cibles sont présents dans les échantillons à des concentrations inférieures à la limite de détection de la méthode GC. Par conséquent, 2 approches différentes pour traiter les zéros ont été évaluées. Dans la première méthode, toutes les impuretés ayant une aire de pic de zéro ont été exclues pendant la comparaison numérique de 2 profils, mais le nombre de pics manquants était calculé pour chaque comparaison.

Dans la deuxième méthode, les zéros ont été remplacés par une valeur correspondant approximativement à la moitié de la limite de détection de la méthode GC (soit 200).

L'effet du prétraitement des données a été évalué de 2 façons. Premièrement, des modèles PLS-DA ont été calculés pour contrôler l'effet du prétraitement sur la capacité à discriminer les échantillons synthétisés par différentes voies. Trois variables « fictives » ont été utilisées dans ces calculs, une pour chaque voie de synthèse (c'est-à-dire les méthodes de Leuckart, de l'amination réductrice et du nitrostyrène).

Le deuxième type d'évaluation de l'impact du prétraitement des données s'est concentré sur la capacité des différentes méthodes à discriminer les échantillons liés et non liés. Dans cette étude, les échantillons liés ont été définis comme étant des échantillons du même lot de fabrication, c'est-à-dire synthétisés de manière hautement reproductible selon la même recette, et les non liés des échantillons issus de différents lots.

De manière concomitante, les 7 métriques ont été utilisées dans l'analyse en PLS-DA. L'importance de certains composés cibles dans la discrimination entre les voies de synthèse et les lots a été évaluée. Le nombre de composés cibles discriminants pour la comparaison d'échantillons d'amphétamine a ainsi été réduit progressivement à 26 au lieu de 33. Enfin, les performances de la méthode de cette nouvelle méthode de profilage ont été comparées à une méthode de référence (méthode du quotient) déjà en place dans certains laboratoires de police scientifique.

Pour réaliser les calculs de comparaison des échantillons par paires, j'ai créé une macro en Visual Basic pour Excel. En annexe 1, le fonctionnement du code est fourni pour la version finale utilisant 2 métriques, Pearson et Canberra et les prétraitements afférents.

Pour les analyses multivariées en PLS-DA, j'ai utilisé le logiciel The Unscrambler® version 7.6 de Camo.

Finalement, la corrélation de Pearson s'est montrée être la métrique qui distinguait le plus efficacement les échantillons liés et non liés. Le remplacement des zéros par 200 et le prétraitement par normalisation à la somme des aires de pic suivie de la racine quatrième se sont avérés être l'alternative de prétraitement la plus efficace. En outre, le nombre de composés cibles utilisés dans la comparaison des échantillons d'amphétamine a pu être réduit à 26 sans perte significative d'informations.

La procédure de profilage mise au point (appelée SMT) s'est révélée supérieure à une méthode de profilage existante qui a été remplacée. La méthode SMT est toujours utilisée actuellement pour le profilage de l'amphétamine.

Suite au projet SMT, la Finlande a été la première à agir en développant une base de données pour stocker les données acquises avec la nouvelle méthode optimisée et permettre des comparaisons de profils dans le cadre de leur programme national de profilage. Il a été développé dans le Crime Laboratory du National Bureau of Investigation à Vantaa, en Finlande (partenaire 2 du projet SMT).

La base de données permettait d'afficher les 50 meilleures correspondances lorsqu'un nouvel échantillon était comparé au reste de la base de données et affichait également des représentations graphiques des chromatogrammes très utiles pour les comparaisons visuelles directes (confirmation de liens par des experts). Cette base de données a été implémentée à l'Institut forensique des Pays-Bas où la nouvelle méthode de profilage de l'amphétamine a été appliquée en routine dans le cadre d'un autre projet de recherche européen qui a reçu l'acronyme CHEDDAR (Collaborative and Harmonized European Database for the Determination of Amphetamine Relations).

Ce projet de 2 ans était financé à travers le programme AGIS (Direction générale de la justice et des affaires intérieures) de la Commission européenne et le projet a

officiellement démarré en décembre 2003 sous la coordination du laboratoire de police scientifique de Suède. Le consortium était composé de la Suède, de la Finlande, de la Grande-Bretagne, de la Pologne, du Danemark et de la France.

J'étais arrivée à cette période au laboratoire de police scientifique de Lyon et ai été impliquée sous cette égide dans ce nouveau projet.

Il était en quelque sorte la continuation du projet avec comme 1^{er} objectif le développement d'une base de données, similaire à la base de données des Finlandais dont le serveur serait physiquement situé dans un laboratoire du consortium et d'autres partenaires pourraient télécharger leurs données directement via Internet. Ils pouvaient également consulter la base de données à partir de leur ordinateur local.

Un 2^e objectif était d'améliorer les connaissances sur les corrélations entre les recettes utilisées dans la fabrication clandestine d'amphétamines et les profils correspondants. Le 3^e objectif était de documenter entièrement la méthode afin de simplifier le processus d'apprentissage pour d'éventuels nouveaux laboratoires qui souhaiteraient appliquer la méthode. Cela comprenait également la conception de macros à des fins de contrôle qualité, la détermination de critères de contrôle qualité pour de nouveaux partenaires potentiels, les problèmes d'archivage et de suivi d'audit, le renvoi d'échantillons de contrôle à des fins de contrôle qualité.

Profilage des impuretés de la MDMA et de la méthamphétamine

Suite aux projets SMT et CHEDDAR, un autre grand chantier a été lancé sous la forme d'un autre projet de recherche financé par la CE dans son 6^e programme-cadre. Il s'agissait du projet CHAMP pour « Collaborative Harmonisation of Methods for Profiling of Amphetamine Type Stimulants ». Les principaux objectifs du projet comprenaient l'harmonisation des méthodes de profilage de la MDMA et de la méthamphétamine (MA) avec la création d'une base de données commune dans une perspective de renseignement sur les drogues.

Le consortium était composé de 8 institutions de police scientifique : Finlande (coordinateur), République tchèque, Pays-Bas, France, États-Unis, Suisse et Allemagne. La durée du projet était de 27 mois entre 2004 et 2006.

Le laboratoire de police scientifique de Lyon était partenaire.

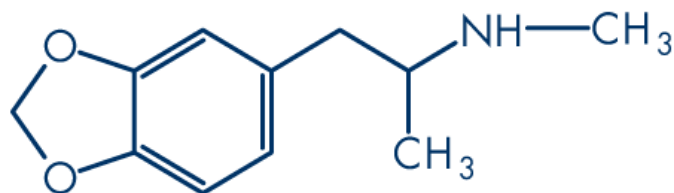
Dans les étapes préliminaires de ce projet, les laboratoires participants ont analysé les caractéristiques physiques, la composition chimique et les impuretés organiques des comprimés de MDMA, en partant de la méthode précédemment harmonisée pour l'amphétamine.

Pour la méthamphétamine, étant présente principalement sous forme de poudre, seules la composition chimique et les impuretés organiques ont été étudiées en partant de la même méthode que celle utilisée pour la MDMA.

Carte d'identité de la MDMA

Nom donné, initialement, à un dérivé amphétaminique, la MDMA (3,4- méthylène-dioxy-méthamphétamine), qui a été synthétisée en 1912. Actuellement, l'ecstasy est un terme générique regroupant aussi d'autres dérivés amphétaminiques dont certains sont classés stupéfiants tels que : DMA, DOET, MMDA, N-hydroxy MDA, MDEA, PMA, STP ou DOM, MDA, TMA, 4-MTA, 2 CI, MBDB, 2 CB, DOB, etc.

Structure moléculaire :



Sa formule chimique est : C₁₁H₁₅NO₂ de masse moléculaire 193.2 g/mol.

Comme avec les autres phénéthylamines, la MDMA existe également sous deux formes énantiomères (R et S).

Le sel le plus courant est le chlorhydrate, présent sous forme de poudre blanche ou blanc cassé ou de cristaux solubles dans l'eau. On rencontre également le sel de phosphate. Les produits illicites se rencontrent essentiellement sous forme de comprimés comportant un logo caractéristique et, moins souvent, sous forme de poudres blanches ou de gélules. La MDMA base est une huile incolore insoluble dans l'eau.

Alors que les phénéthylamines non substituées sur le cycle exercent généralement des effets comparables à ceux des stimulants, la substitution sur le cycle (comme dans le cas de la MDMA) induit une modification des propriétés pharmacologiques. L'ingestion de MDMA provoque de l'euphorie, renforce la conscience sensorielle et induit une légère stimulation du système nerveux central. Elle est moins hallucinogène que son homologue inférieur, la méthylènedioxyamphétamine (MDA). Les termes empathogène et entactogène ont été créés pour décrire les effets de socialisation de la MDMA. Sous forme de comprimé, elle est presque systématiquement utilisée par voie orale (ingérée), mais sous la forme de poudre elle peut être sniffée, inhalée, voire ingérée ou injectée bien que la pratique d'injection soit très rarement observée dans les contextes d'usage festifs^{68,69}.

Petite histoire de la MDMA

Appelée aussi « pilule de l'amour », la MDMA est devenue l'emblème de la culture techno dans les années 1980, qui connaît depuis 2010 un renouveau explosif, hors du cercle restreint des clubbeurs et des ravers.

La MDMA a été brevetée en 1913 par la société chimique allemande Merck. Elle était destinée à être vendue comme pilule amaigrissante, bien que le brevet ne mentionne aucune utilisation spécifique. La société a décidé de ne pas commercialiser le médicament. L'armée américaine a expérimenté la MDMA en 1953 lors de recherches sur la guerre psychologique.

Le chimiste qui popularisa la MDMA en 1976 est Alexander Shulgin. Il dédia sa vie à la synthèse de substances psychoactives et préconisait un usage thérapeutique de la MDMA. Il a notamment coécrit avec Ann Shulgin 2 livres : *Phenethylamines I Have Known And Loved: (PiHKAL): À Chemical Love Story* publié en 1991 (<http://isomerdesign.com/PiHKAL/index.php?domain=pk>) et *TiHKAL: The Continuation*" (<http://isomerdesign.com/PiHKAL/?domain=tk>).

⁶⁹ https://www.emcdda.europa.eu/publications/drug-profiles/mdma_fr

Quatre précurseurs principaux peuvent être utilisés dans le cadre de la fabrication de la MDMA et de drogues apparentées : le **safrole**, l'**isosafrole**, le **pipéronal** et la 3,4-méthylènedioxyphényl-2-propanone (**PMK**). Le safrole est le matériau initial de base, les 3 autres substances pouvant être synthétisées à partir de ce dernier. Dans le brevet original de Merck obtenu en 1914, le safrole a été mis à réagir avec de l'acide bromhydrique pour former du bromosafrole, qui a ensuite été converti en MDMA en utilisant de la méthylamine. De nombreuses synthèses illicites démarrent avec de la PMK et utilisent soit la synthèse de **Leuckart**, soit diverses **aminations réductrices**, y compris le procédé de la feuille d'aluminium. Tous ces procédés donnent de la MDMA racémique ⁶⁷. La figure 33 montre les différentes voies de synthèse issues du safrole (figure d'après ⁷⁰) :

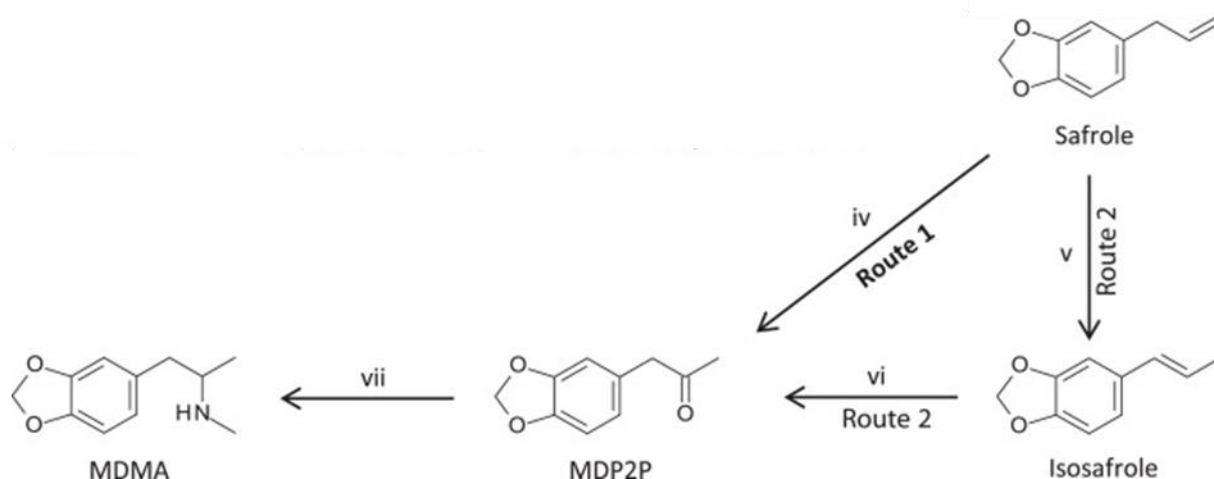


Figure 33 : Synthèse de la MDMA à partir du safrole : (iv) *p*-benzoquinone, PdCl₂ ; (v) KOH ; (vi) 1: H₂O/HCOOH, 2: H₂SO₄ ; (vii) CH₃NO₂, Al/Hg.

Le profil chimique de la MDMA est fortement affecté par différents paramètres de fabrication comme la voie de synthèse, les produits chimiques (solvants, etc.) ou la qualité des précurseurs (il n'est pas rare non plus que différents lots synthétisés soient mélangés avant le processus de mise en comprimé). Ces paramètres, et le fait que les processus de cristallisation, de séchage et d'ensemencement ne soient pas effectués de manière contrôlée contribuent à la grande variabilité du produit final. Dans cette recherche, il a été considéré que le mélange final préparé pour la fabrication des comprimés avait un profil chimique organique spécifique, défini comme le lot de précompression (pre-TB pour pre-tabletting).

La poudre mélangée est ensuite comprimée à l'aide d'une machine à comprimer (presse) dotée de poinçons interchangeables afin d'imprimer un logo et/ou une marque sur la surface des comprimés (comme un champignon ou un fantôme ce qui leur donne l'apparence d'un bonbon et peut servir aussi d'objet de marketing auprès des consommateurs). La forme du poinçon est généralement un paramètre fixe pour une machine à comprimer donnée, tandis que l'épaisseur dépend de la quantité de poudre par comprimé et du réglage de la pression des poinçons. Une fois pressé, le comprimé ne peut plus être altéré et il conserve ses caractéristiques tout au long des étapes du trafic jusqu'à ce qu'il soit consommé ou saisi par les services de police. Un lot de fabrication de comprimés (post-TB) est défini comme étant produit par une compresseuse spécifique avec des réglages donnés (cf. figure 34 issue de ⁷¹).



Figure 34 : Exemples de logo de comprimés d'ecstasy

Les différentes étapes de la production (pré-B et post-TB) peuvent être réalisées dans des endroits différents et donner lieu à des ensembles indépendants de caractéristiques. C'est pourquoi ces caractéristiques sont analysées séparément afin d'étudier leur potentiel d'individualisation des échantillons provenant d'un même lot de fabrication pre-TB (1^{re} partie) ou d'un même lot post-TB (2^e partie). Une fois combinées, ces deux sources d'information apportent une valeur ajoutée au processus de comparaison.

Dans cette étude, les échantillons de MDMA provenaient de cas réels (saisies policières). Les profils des comprimés de chaque saisie ont été considérés comme étant homogènes. Des échantillons tests étaient constitués de 26 saisies provenant de Finlande et de France. Pour chacun des 26 échantillons, chaque partenaire a reçu 10g de poudre homogénéisée (provenant de comprimés broyés). Les analyses répétées 3 fois, effectuées sur les échantillons de test dans 4 laboratoires différents, ont constitué la population d'échantillons avec des profils d'un même lot pre-TB. Quarante-vingt échantillons de rue ont été collectés par chaque laboratoire parmi les saisies effectuées dans leur propre pays entre 1996 et 2004 (Finlande, Pays-Bas, France et Suisse) et ont constitué la population d'échantillons (c'est-à-dire 320 échantillons) avec des profils de lots pré-TB différents.

J'ai travaillé sur le profilage des impuretés de la MDMA qui a consisté à sélectionner les composés cibles. Au départ, ils étaient 46 et ont été ensuite réduits à 32. Finalement, 2 métriques, le coefficient de corrélation de Pearson et la fonction cosinus, appliquées sur des données prétraitées avec la racine carrée ont fourni une très bonne discrimination entre les différents échantillons.

Des analyses en composantes principales (ACP) ont également été entreprises pour visualiser des trafics clandestins et les relier à des impuretés en particulier comme le montre la figure ci-après :

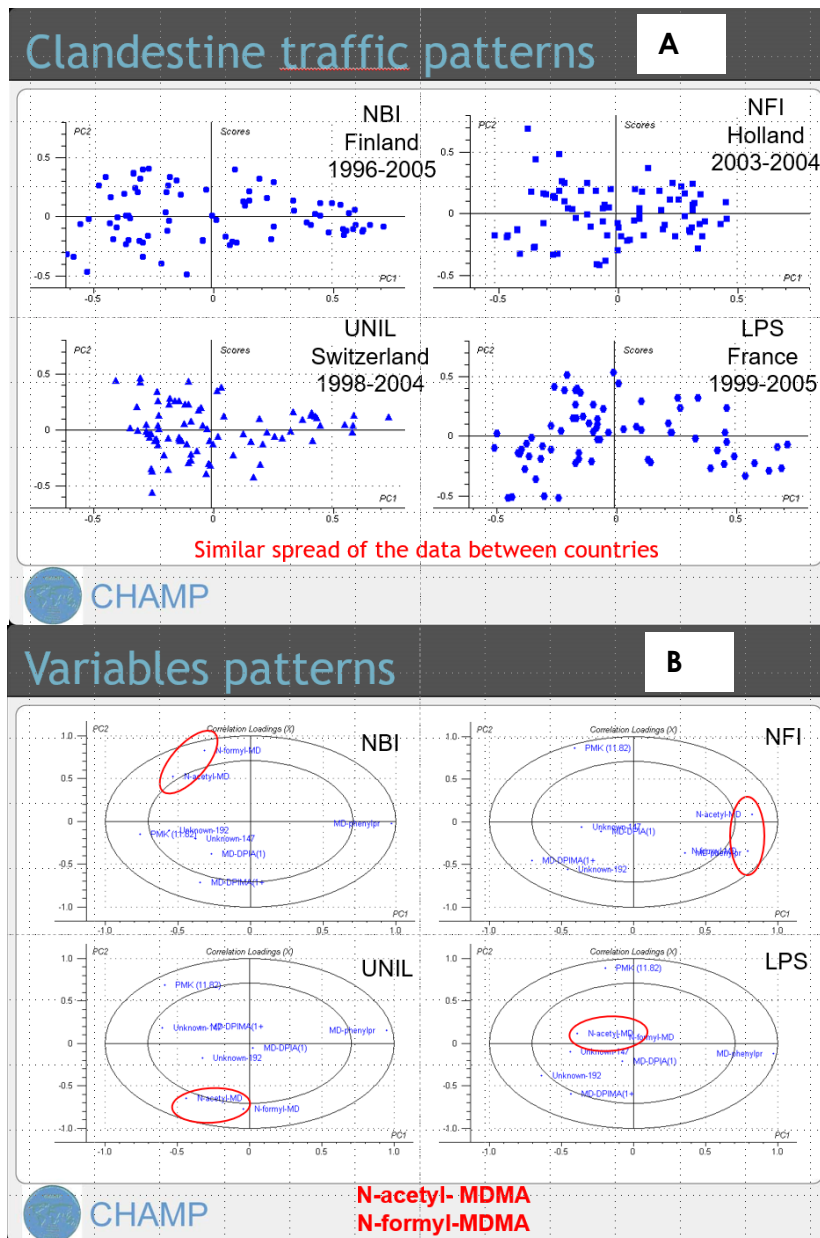


Figure 35 : ACP d'échantillons de rue provenant de différents pays, analysés par la méthode harmonisée. (A) graphique des individus (scores plot) et (B) graphiques des variables.

Les résultats de cette tâche du Work package (WP) 4 « Profiling of Amphetamine Type Stimulant samples » ont été publiés dans la revue Journal of Forensic Science ⁷². Une autre tâche a porté sur l'étude des caractéristiques macroscopiques des comprimés (post-TB). Des analyses multivariées et des comparaisons de prétraitement + métriques sur une dizaine de variables ont été réalisées. Les résultats ont fait l'objet d'un autre article ⁷².

Profilage de la méthamphétamine

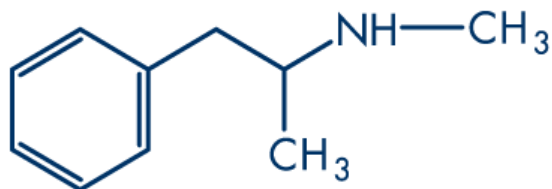
L'étude du profilage des impuretés organiques de la méthamphétamine faisait également partie des tâches du WP4 du projet.

La table des matières du rapport final du projet, qui est confidentiel, figure en annexe 2.

Carte d'identité de la méthamphétamine

La méthamphétamine est la N,-diméthylphénéthylamine.

Structure moléculaire :



Sa formule moléculaire est C₁₀H₁₅N et son poids moléculaire : 149.2 g/mol.

L'atome de carbone asymétrique donne lieu à deux énantiomères, stéréo-isomères R et S.

La méthamphétamine base est une huile volatile incolore insoluble dans l'eau. Le sel le plus courant est le chlorhydrate. Il existe sous forme de poudre blanche ou blanc cassée ou de cristaux solubles dans l'eau. Les produits illicites sont généralement constitués de poudres, mais le chlorhydrate cristallin pur est connu sous le nom de « Ice ». Les comprimés contenant de la méthamphétamine peuvent comporter des logos similaires à ceux figurant sur les comprimés de MDMA et autres comprimés d'ecstasy.

La méthamphétamine est un stimulant du SNC qui induit une hypertension et une tachycardie ainsi qu'une plus grande confiance en soi et une sociabilité et une énergie accrues. Elle inhibe l'appétit, supprime la fatigue et provoque de l'insomnie. Les effets apparaissent généralement dans les 30 minutes qui suivent la prise orale et perdurent pendant de nombreuses heures. La méthamphétamine est plus puissante que l'amphétamine, mais dans les situations non contrôlées, les effets sont quasiment indifférenciables^{p 73}.

^p https://www.emcdda.europa.eu/publications/drug-profiles/methamphetamine_fr

Petite histoire de la méthamphétamine

La méthamphétamine a été développée au Japon en 1919. Elle a été largement utilisée durant la Seconde Guerre mondiale, quand des deux côtés on l'employait pour garder les troupes éveillées. De fortes doses furent données aux pilotes kamikazes japonais avant leurs missions suicides, et après la guerre, l'abus de méthamphétamine en intraveineuse atteignit des sommets quand les réserves militaires ont été mises à la disposition du public japonais.

Dans les années 1950, la méthamphétamine était prescrite comme soutien à un régime alimentaire et pour lutter contre la dépression. D'accès facile, elle était utilisée comme stimulant non médical par les étudiants, les chauffeurs routiers, les athlètes notamment. La situation a nettement changé dans les années 1960. La méthamphétamine injectable est devenue largement disponible et la consommation a fortement augmenté. Puis, en 1970, le gouvernement américain en a fait une drogue illégale. Les gangs de motards contrôlaient alors la plus grande partie de la production et de la distribution de cette drogue. La plupart des consommateurs de l'époque vivaient dans des régions rurales et ne pouvaient pas se payer de la cocaïne, qui était plus chère. Dans les années 1990, les organisations mexicaines qui se livraient au trafic de drogue ont installé de grands laboratoires en Californie. Alors que ces laboratoires pouvaient fabriquer 25 kilos de substance en un seul week-end, de petits laboratoires privés ont éclos dans les cuisines et les appartements, ce qui a valu à cette drogue le nom de « stove top » (dessus de fourneau). C'est depuis cette époque qu'elle s'est répandue aux États-Unis et en Europe, via la République tchèque. Aujourd'hui, la plus grande partie de la drogue disponible en Asie est produite en Thaïlande, au Myanmar et en Chine.

L'énantiomère S est généralement produit par la réduction de la l-**éphédrine** ou de la d-**pseudoéphédrine**. Il est en proportion supérieure à l'énantiomère R. L'éphédrine et la pseudoéphédrine sont toutes 2 disponibles dans le commerce et sont utilisées dans certains médicaments. L'éphédrine peut également être extraite de la plante ***Ephedra vulgaris L.*** (connue dans la médecine chinoise sous le nom de Ma Huang). Les synthèses de **Leuckart** et de **l'amination réductrice** (procédé de la feuille d'aluminium par exemple) de la 1-phényl-2-propanone (**P2P**, **BMK**, phénylacétone) produisent un mélange racémique des énantiomères R et S.

Ci-après (figure 36) les principales voies de synthèse utilisées dans la fabrication illicite de méthamphétamine d'après ^{67,74} :

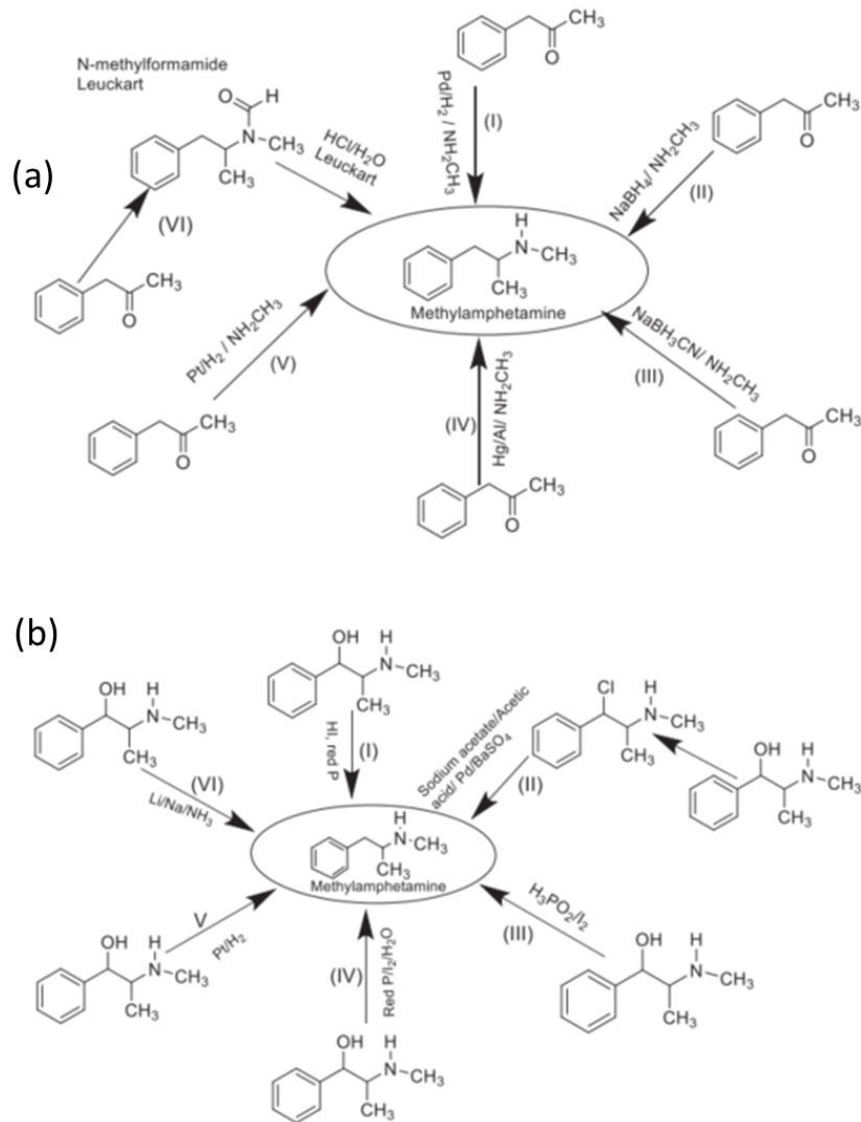


Figure 36 : Voies de synthèse : (a) à partir de la BMK (voie de Leuckart) : (I) $\text{Pd}/\text{H}_2/\text{NH}_2\text{CH}_3$; (II) $\text{NaBH}_4/\text{NH}_2\text{CH}_3$; (III) $\text{NaBH}_3\text{CN}/\text{NH}_2\text{CH}_3$; (IV) $\text{Hg}/\text{Al}/\text{NH}_2\text{CH}_3$; (V) $\text{Pt}/\text{H}_2/\text{NH}_2\text{CH}_3$; (VI) $\text{HCl}/\text{H}_2\text{O}$ et (b) à partir de éphédrine/pseudoéphédrine : voie (I) Nagai ; (II) Emde ; (III) ; Hypo ; (IV) Moscow ; (V) Rosenmund ; (VI) Nazi/Birch.

Dans le projet CHAMP, des échantillons de méthamphétamine (MA) provenant des 2 voies de synthèse de la BMK et de l'éphédrine/pseudoéphédrine (EPH) ont été collectés.

Pour garantir la comparabilité des résultats des différents laboratoires utilisant la même méthode de profilage, des échantillons de contrôle ont été utilisés et les résultats ont été validés par un essai en aveugle interlaboratoire.

Les échantillons provenaient de 27 saisies provenant des États-Unis, de la République tchèque et d'Australie. Des répliquas, 6 par laboratoire et par échantillon de contrôle et 3 échantillons de contrôle ont été utilisés. Vingt-quatre échantillons de test, 1 répliqua par laboratoire et par échantillon de test, ont été analysés dans 4 laboratoires différents et ont constitué la population des échantillons liés.

En outre, les échantillons de rue collectés par 2 laboratoires (Finlande et République tchèque) à partir de saisies locales ont constitué la population d'échantillons non liés : au total, 151 échantillons : 88 échantillons de Finlande, 4 d'Estonie, 2 de Norvège, 15

du Danemark et 42 de la République tchèque. Trois répliquas de chaque échantillon de rue ont été analysés.

Tous les partenaires disposaient du même équipement de chromatographie (GC-MS). Une analyse statistique univariée a permis une 1^{re} sélection de 43 composés cibles (sous-produits formés pendant la réaction) : le coefficient de variation (CV ou relative standard deviation RSD) a été calculé pour chacune des variables entre les laboratoires afin de sélectionner le premier ensemble de composés cibles sur la base de la reproductibilité. L'opération a ensuite été répétée pour les échantillons de rue afin d'évaluer l'étendue de la variation de chaque variable dans l'ensemble d'échantillons non liés et de réaliser une deuxième sélection de variables. La fréquence d'apparition des variables a été calculée en enregistrant leur présence ou leur absence dans les échantillons de rue. Une étude de la corrélation entre les composés cibles a ensuite été entreprise afin de réduire le nombre de variables. La sélection finale des composés cibles, à ce niveau d'analyse, a consisté en une combinaison de critères de corrélation, de stabilité et de qualité chromatographique (pic bien intégré). Il en est resté 24.

Ensuite, des analyses statistiques multivariées ont été entreprises. Plusieurs prétraitements et métriques ont été testés afin de déterminer la combinaison permettant la meilleure discrimination entre des échantillons provenant de différents lots de fabrication. Pour ce faire, des macros VBA pour Excel ont été conçues, une pour réaliser, tous les prétraitements choisis et l'autre toutes les métriques y compris la prise en compte des pics manquants. Une copie de la feuille de démarrage est montrée en figure 37 :

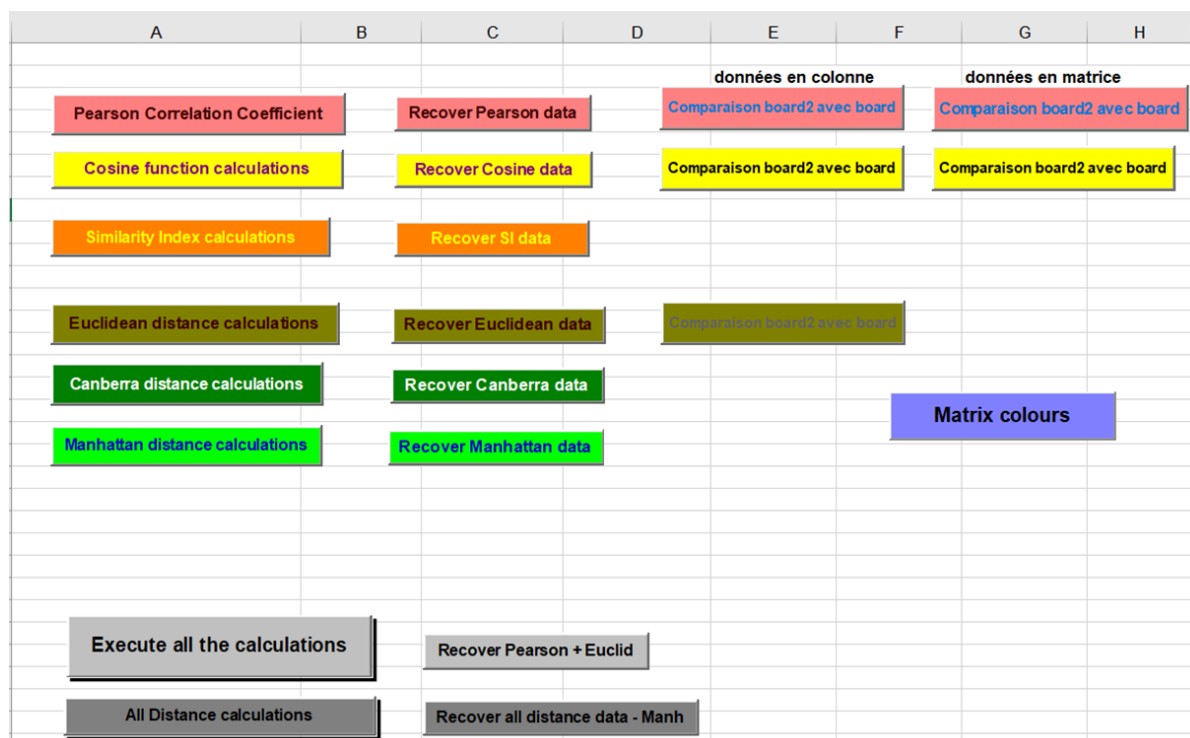


Figure 37 : Copie d'écran de la feuille de démarrage de la macro pour les calculs des différentes métriques.

Afin de pouvoir comparer les métriques entre elles, elles ont été recalculées sur une même échelle qui pouvait être entre 0 et 100, avec 100 représentant le « match » parfait et 0 la différence totale ou l'inverse. Le maximum pour les mesures de distances était donc calculé en amont pour chaque jeu de données à comparer.

Deux formats de sortie de données étaient disponibles, soit toutes les valeurs en colonne pour pouvoir ensuite visualiser la distribution de celles-ci, soit sous forme de matrice dont une moitié fournissait le nombre de composés cibles pris en considération dans les échantillons comparés et l'autre moitié les valeurs de distance avec possibilité de colorer certaines valeurs au-dessus d'un certain seuil. Un exemple est fourni dans la figure 38 :

	A	B	C	D	E	F	G	H	I	J	K	L
1			2013/001843	2013/008671	2013/012026	2013/015629	2013/024181	2013/026925	2013/027658	2013/027811	2013/028033	2013/030744
2	2013/001843	5	100	12.44834463	25.94780526	10.28066266	7.165373295	15.78905918	66.44668051	60.41966988	21.86042114	19.34868908
3	2013/008671	5	5	100	96.04848192	99.86301018	99.11213202	99.34610919	66.56791738	73.17019279	98.19588257	98.68029285
4	2013/012026	5	5	5	100	94.56735721	92.65600032	97.4116498	82.83197973	87.71229112	99.51812276	99.15648593
5	2013/015629	5	5	5	5	100	99.55944735	98.97098916	63.13306849	69.94283198	97.11775564	97.77510913
6	2013/024181	5	5	5	5	5	100	97.86070252	58.51208744	65.38743853	95.35374829	96.42112716
7	2013/026925	5	5	5	5	5	5	100	70.77472386	77.11739401	98.91828652	99.10211626
8	2013/027658	5	5	5	5	5	5	5	100	99.44938669	78.39930029	75.97591459
9	2013/027811	5	5	5	5	5	5	5	5	100	83.96221596	81.72254284
10	2013/028033	5	5	5	5	5	5	5	5	5	100	99.81712533
11	2013/030744	5	5	5	5	5	5	5	5	5	5	100
12	2013/032084	5	5	5	5	5	5	5	5	5	5	5

Figure 38 : Exemple de sortie sous forme de matrice indiquant, d'une part, le nombre de composés cibles, ici 5, et d'autre part, les valeurs de la métrique entre les échantillons, ici la fonction cosinus ramenée sur une échelle entre 0 et 100. 100 indique la valeur de la métrique pour la comparaison de l'échantillon avec lui-même. Un seuil de 99.8 a été appliqué ici et toutes les valeurs au-dessus sont matérialisées en gras et en rouge.

La capacité à discriminer les échantillons entre eux a été évaluée pour chaque combinaison prétraitement + métrique en utilisant le rapport DP (discriminant power) calculé selon l'équation 3 avec $k = 3$ dans ce cas.

La meilleure discrimination a été obtenue avec la combinaison de prétraitement normalisation suivie de la racine carrée ($N+2R$) et l'application de la fonction cosinus carré sur les données prétraitées. Une autre option correspondant à la combinaison ($N + 2R$) et à l'application de coefficient de Pearson modifié ($\frac{1-\rho}{2} \times 100$) sur les données prétraitées a montré des DP élevés. Cependant, elle fournissait un nombre de faux négatifs moins élevé.

Des analyses multivariées ont été entreprises afin d'évaluer la possibilité, d'une part, de séparer les échantillons provenant des 2 principales voies de synthèse, via la BMK, occasionnellement utilisée en Europe et aux États-Unis et via l'EPH, communément utilisée en Asie, Australie et États-Unis, à cette époque-là (années 2000) et d'autre part, de différencier les profils d'impuretés provenant de recettes différentes utilisant la synthèse via l'EPH.

Dans un premier temps, l'ACP a été utilisée sur 3 échantillons de contrôle x 6 répliques (18 objets), 22 échantillons test analysés par 4 laboratoires (88 objets) et 152 échantillons de rue. Pour la majorité, la voie de synthèse certaine était l'EPH, pour certains échantillons, la voie de la BMK était certaine, d'autres elle était supposée et pour 56, la voie était inconnue.

Les graphiques des individus et des variables sont fournis sur la figure 39. Le pourcentage de variance expliquée avec 3 composantes est de 67%, ce qui est satisfaisant compte tenu de la nature et de la provenance des échantillons.

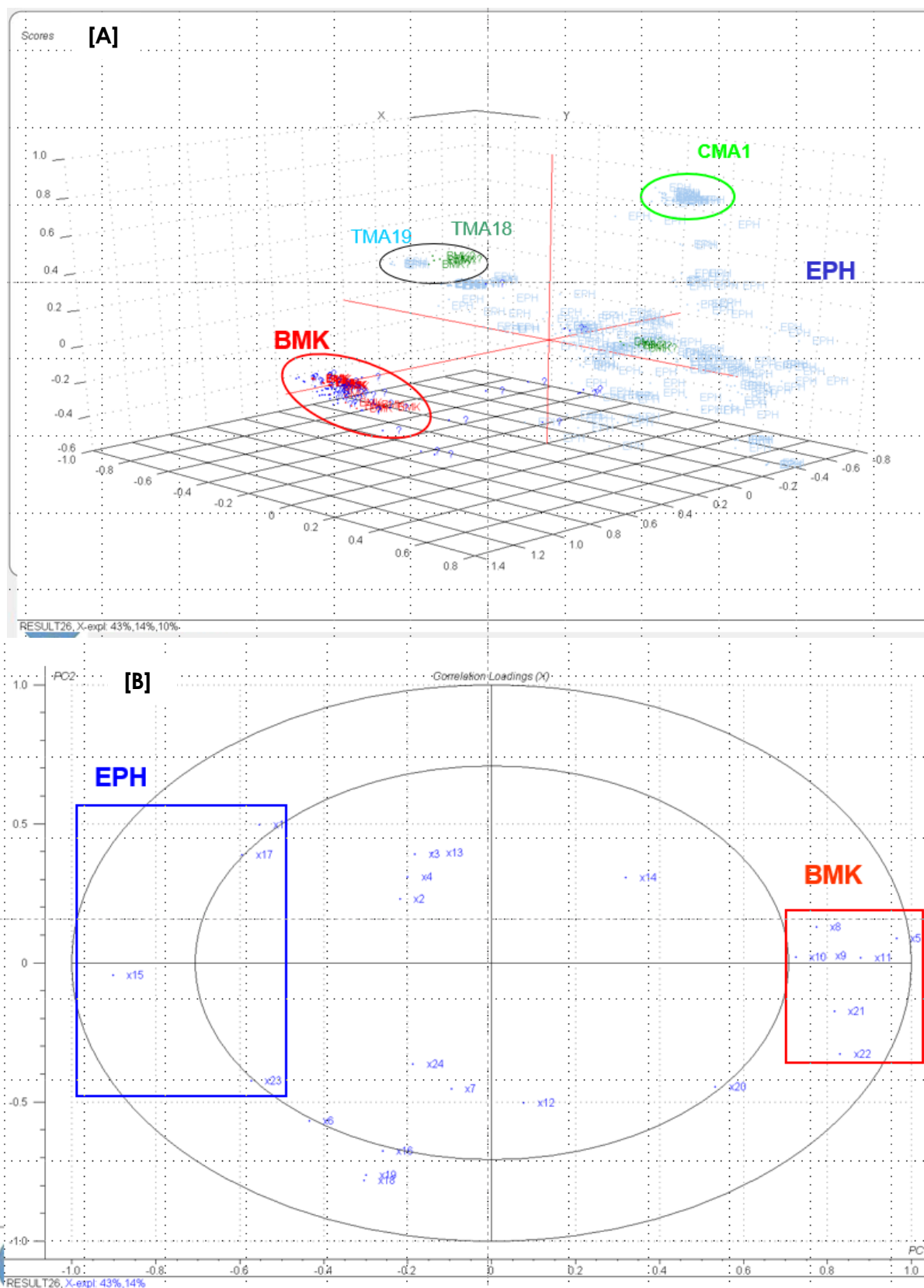


Figure 39 : ACP d'échantillons de MA issus des contrôles, des tests et de rue. [A] graphique des individus. [B] graphiques de variables.

Les points d'interrogation correspondent aux échantillons dont la voie est incertaine, on peut voir qu'ils correspondent pour la plupart à la voie de la BMK. Sur le graphique

des variables, on distingue nettement quelles sont les variables permettant de différencier la voie de synthèse. Pour l'EPH, il s'agit du composé x15 qui n'avait pas pu être identifié à l'époque et nommé Unknown 58d et pour la BMK, c'est essentiellement la para-t-butylmethamphetamine, composé x5.

Sur 90 échantillons de rue issus de la voie de l'EPH provenant de République tchèque et de Finlande, une classification hiérarchique ascendante (CAH) a été appliquée pour étudier la possibilité de regrouper les échantillons en fonction d'une recette en particulier. Plusieurs méthodes ont été testées et c'est la méthode utilisant la distance euclidienne et l'agrégation de Ward qui a donné les meilleurs résultats. On a obtenu 6 groupes comme le montre le dendrogramme figure 40 :

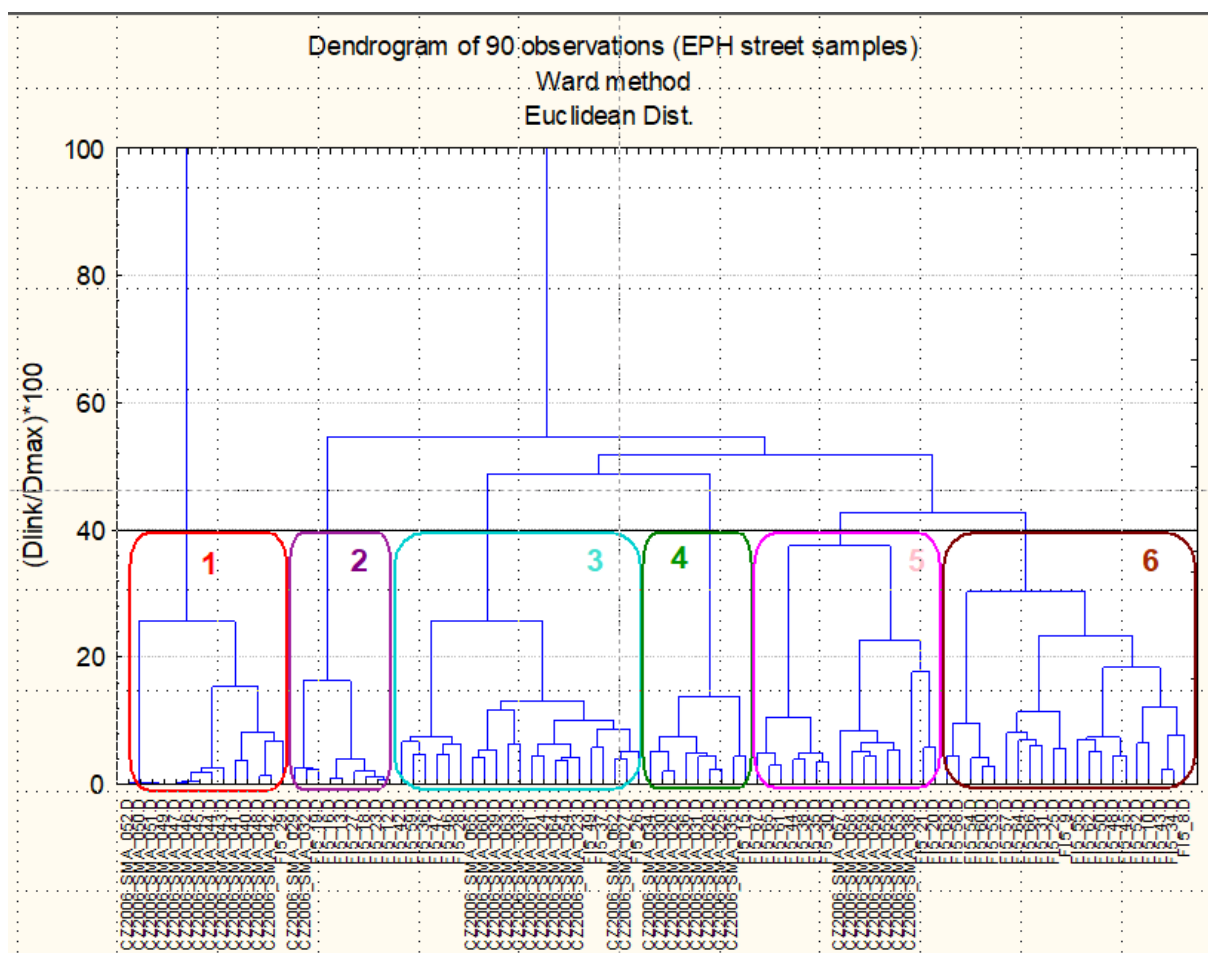


Figure 40 : Dendrogramme de la CAH de 90 échantillons de rue provenant de République tchèque et de Finlande.

Après analyse des profils d'impuretés des échantillons, dans la classe 1 on retrouve des échantillons avec une présence élevée de pseudoéphédrine, la classe 2 est caractéristique de composés de la famille des naphtalènes indiquant que l'éphédrine a été réduite dans des conditions acides via le mélange HI/Phosphore rouge typique de la recette de Nagai. Dans la classe 3, on retrouve des échantillons avec un composé majoritaire qui est l'aziridine. La classe 4 comprend des échantillons dont l'éphédrine est le composé majoritaire. La classe 5 est caractéristique d'échantillons avec de la N-formylméthamphétamine et de la benzoylméthamphétamine et enfin la classe 6 comprend des échantillons avec majoritairement de la méthamphétamine dimère. Ces résultats étaient prometteurs, car ils permettaient de séparer les

échantillons de différentes voies de synthèse, mais pour étudier plus en détail à quoi correspondaient tous ces composés cibles, il aurait fallu pouvoir synthétiser en laboratoire des échantillons suivant toutes les recettes possibles, ce qui n'était pas envisageable pendant la durée du projet.

Enfin, une PLS-DA (analyse discriminante par les moindres carrés partiels) sur 324 échantillons de synthèse connue a été utilisée pour classer les échantillons provenant des voies BMK et EPH. Les données ont été, au préalable, prétraitées avec une normalisation suivie d'une racine quatrième pour réduire la dispersion interlaboratoires. Pour valider le modèle, une validation croisée de type « leave-one-out » a été utilisée suivie d'un test à l'aveugle avec 64 échantillons. Seuls 3 échantillons ont été mal classés représentant un taux d'erreur de 5%.

La plupart des résultats ont été publiés dans la revue *Forensic science International* ⁷⁵.

L'un des principaux changements observés ces dernières années (à partir de 2017) a été l'augmentation de la pureté de la méthamphétamine, au point qu'une grande partie de la méthamphétamine saisie peut être décrite comme étant presque pure. Elle est dite "de qualité pharmaceutique" et est si pure qu'elle contient peu ou pas de sous-produits de fabrication, utiles pour déterminer la voie de synthèse et les précurseurs chimiques ayant servi à la fabrication. L'analyse de la méthamphétamine par les techniques classiques de profilage chimique, c'est-à-dire le profilage des impuretés organiques par chromatographie en phase gazeuse/spectrométrie de masse, l'analyse chirale et l'analyse élémentaire, ne fournit alors que peu d'informations sur le ou les précurseur(s) et le procédé de synthèse utilisés pour fabriquer les échantillons ⁷⁶.

Le profilage chimique non ciblé a été examiné pour déterminer son potentiel sur ce type d'échantillon.

Dans le profilage chimique non ciblé, l'objectif est d'identifier et de quantifier tous les composés de l'échantillon ou du moins, le plus grand nombre possible. Cette méthode s'oppose à l'analyse ciblée plus traditionnelle où l'expérience est optimisée pour quantifier un ou quelques composés. Dans cette optique, j'ai co-encadré une thèse dont l'objectif était de développer une méthode pour détecter et identifier des impuretés pour le profilage chimique complémentaire de la méthamphétamine ⁷⁷. Les signatures chimiques ont été caractérisées à l'aide d'une chromatographie liquide à ultra-haute performance couplée à une spectrométrie de masse en tandem avec ionisation par électrobulbion (UHPLC-MS/MS).

La méthode mise au point a permis de détecter et d'identifier plus de 90 impuretés. La méthamphétamine saisie et analysée avait une pureté moyenne de 95 % en principe actif. Les profils basés sur les impuretés co-extraites comme la noréphédrine et la méthyléphédrine ont été examinés et il a été découvert que cette dernière se dégrade pour former de la N, N-diméthylamphétamine, également détectée dans les échantillons analysés. Les profils de « co-extraction » ont été utilisés pour confirmer la provenance de l'éphédrine de la plante éphédra.

Les impuretés ont été semi-quantifiées et des impuretés telles que la N, N-diméthylphénéthylamine, la phentermine, l' α -éthylphénéthylamine, la N-éthylphénéthylamine et la N, N,-4-triméthylphénéthylamine, ont été retrouvées à des concentrations très faibles (moins de 0,0001 % en poids). Bien que le précurseur soit dérivé de l'éphédra, la possibilité d'une extraction à partir de médicaments en vente libre tels que l'asthmacon (aminophylline/guaifénésine/éphédrine/phénobarbital

systémique), le Mudrane GG (aminophylline/phénobarbital/éphédrine/iodure de potassium systémique), la Quelidrine (carbétapentane/éphédrine/chlorphéniramine/phényléphrine systémique) n'a pu être exclue.

De plus, des composés inconnus ont également été caractérisés structurellement. Leur description a fait l'objet d'un article dans *Forensic Science International: Synergy* ⁷⁴.

Profilage des impuretés de l'héroïne et de la cocaïne

À] Héroïne

Comme mentionné dans la genèse du profilage, c'est l'héroïne qui a été le premier objet d'étude. Très tôt des méthodes analytiques ont été mises au point pour déterminer des impuretés d'intérêt suivant le niveau d'information souhaité. Pour pouvoir mettre au point une méthode de profilage, on l'a vu, il est indispensable de connaître comment est fabriqué le produit.

Carte d'identité de l'héroïne

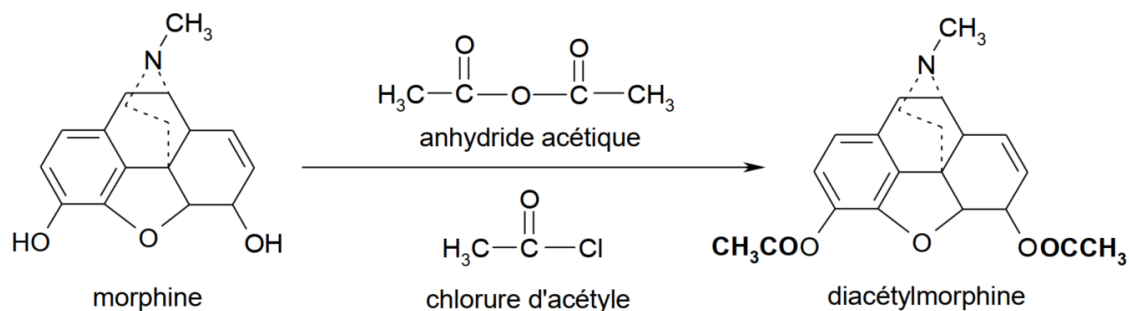
L'héroïne ou diacétylmorphine ou diamorphine est un produit semi-synthétique obtenu par acétylation de la morphine, un produit naturel présent dans l'opium, à savoir le latex séché de certaines espèces de pavot (ex. *Papaver somniferum* L. figure 41).



Figure 41 : Pavot somnifère ou pavot à opium.

Structure moléculaire :

Acétylation de l'héroïne à partir de la morphine



Formule moléculaire de l'héroïne : $C_{21}H_{23}NO_5$ - Poids moléculaire: 369.4 g/mol

Sa fabrication s'opère en plusieurs étapes et nécessite l'utilisation de différents produits chimiques tels que de l'hydroxyde de sodium, du carbonate de calcium ou de sodium, du chlorure d'aluminium, de l'anhydride acétique qui laissent toujours des traces dans le produit final.

Le schéma ci-après (figure 42) décrit ces différentes étapes :

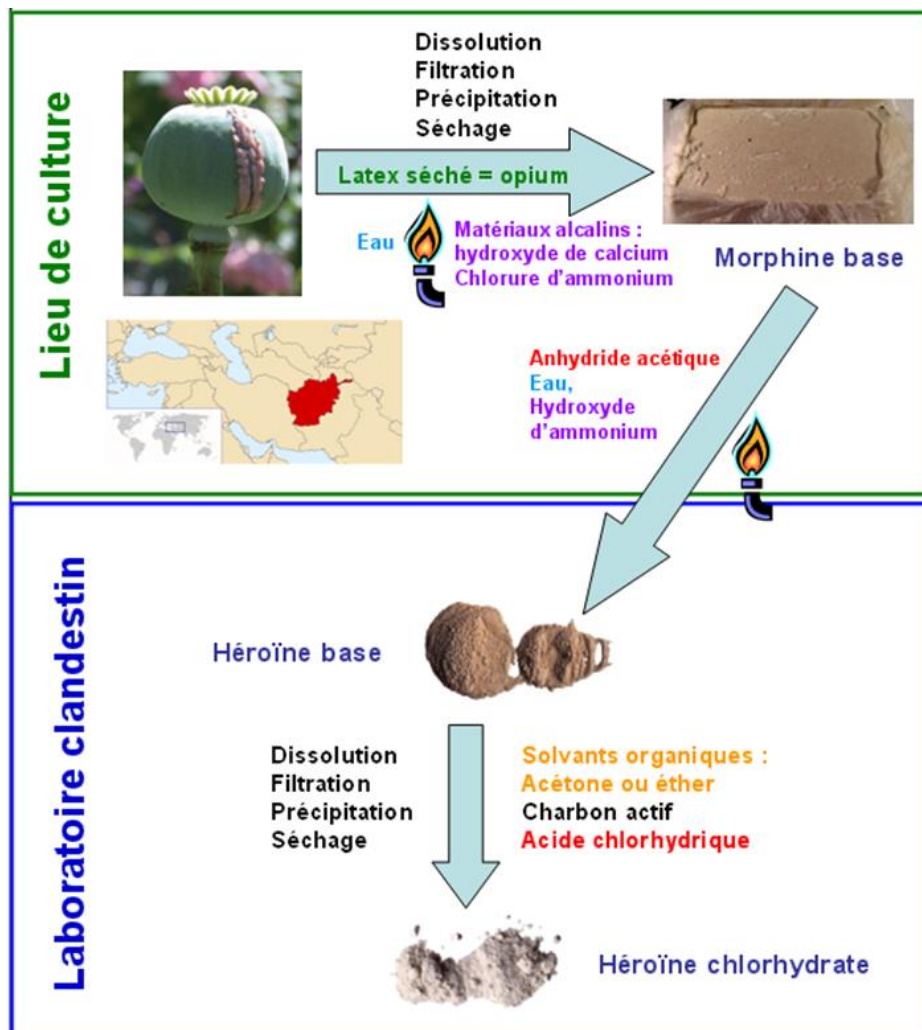


Figure 42 : Étapes de fabrication de l'héroïne à partir du pavot (*Papaver somniferum* L.).

Le latex obtenu à partir des capsules ou « poppies » est mis à sécher. Ce matériau (opium) est dispersé dans une solution aqueuse d'hydroxyde de calcium (chaux éteinte). L'alcalinité est ajustée en ajoutant du chlorure d'ammonium, provoquant ainsi une précipitation de la morphine-base. La morphine séparée est bouillie avec de l'anhydride acétique. L'ajout de carbonate de sodium induit la séparation de la diamorphine base brute. Selon les régions, cette substance peut être consommée telle quelle, purifiée ou transformée en chlorhydrate.

Jusqu'à la fin des années 70, la quasi-totalité de l'héroïne consommée en Europe provenait d'Asie du Sud-Est (Triangle d'Or). Aujourd'hui, l'héroïne est essentiellement issue d'Asie du Sud-Ouest, principalement d'Afghanistan et du Pakistan (Croissant

d'Or). L'héroïne est également produite dans certaines régions d'Amérique du Sud, mais celle-ci est rarement consommée en Europe. L'anhydride acétique, un précurseur essentiel dans le cadre de la fabrication de l'héroïne, est répertorié au Tableau I de la Convention de 1988 des Nations unies contre le trafic illicite des stupéfiants et substances psychotropes^q.

L'héroïne est un opioïde (issu de la semi-synthèse du pavot), dépresseur qui ralentit l'activité du système nerveux central, ses propriétés analgésiques sont beaucoup plus puissantes que les opiacés (opium, morphine, codéine qui sont des dérivés naturels du pavot)^r. Une consommation répétée de diacétyle entraîne une tolérance et une dépendance physique. Elle est 2 à 3 fois plus puissante que la morphine. La dose létale minimale est estimée à 200 mg, mais les personnes dépendantes peuvent tolérer jusqu'à 10 fois cette dose.

L'héroïne disponible en Asie du Sud-Ouest est une poudre brune qui se présente généralement sous la forme de base libre, insoluble dans l'eau, mais soluble dans les solvants organiques. Elle peut être "fumée" en chauffant la forme solide sur une feuille de métal au-dessus d'une petite flamme et en inhalant ensuite la vapeur qui s'en dégage (s'appelle « chasser le dragon »). Pour être injectée, cette forme chimique d'héroïne doit d'abord être solubilisée, par exemple avec de l'acide citrique ou de l'acide ascorbique. L'héroïne base (poudre brune) est la forme chimique la plus présente sur le marché européen notamment français et suisse.

L'héroïne d'Asie du Sud-Est, moins courante en Europe, est généralement une poudre blanche qui se présente sous la forme de sel de chlorhydrate hydraté. Elle est soluble dans l'eau, mais insoluble dans les solvants organiques. Elle peut être injectée directement sous forme de solution.

^q https://www.emcdda.europa.eu/publications/drug-profiles/heroin_fr

^r <https://www.ofdt.fr/produits-et-addictions/de-z/heroine-et-autres-opiacés>

Petite histoire de l'héroïne

Le pavot à opium est connu depuis des milliers d'années, des capsules et des graines ont été retrouvées dans des habitats néolithiques européens datant de 5000 ans av. J.-C.. À la fin du IV^e millénaire av. J.-C., les Sumériens faisaient le commerce de l'opium et de ses graines à travers tout le bassin méditerranéen et jusqu'en Inde. Il était aussi connu des Égyptiens et notamment des pharaons qui le consommaient certes pour ces vertus thérapeutiques, mais aussi pour ces effets psychotropes... Dans la Grèce Antique, Homère, dans l'Odyssée, cite un breuvage à base de vin et d'opium, désigné sous le nom de Nepenthes, boisson donnée aux guerriers pour atténuer la sensation du danger. L'opium fut beaucoup utilisé à l'époque romaine et en particulier à l'époque où règne l'empereur Néron. Andromaque, médecin de la cour, a composé une mixture qui devint célèbre : la thériaque (64 ingrédients, dont le venin de vipère, l'opium, la gentiane, le gingembre, etc.). En Europe, on voit apparaître l'opium plutôt vers le 15^e siècle avec le développement du breuvage du Docteur Paracelse. Boisson ressemblant étrangement à la thériaque, il la nomma le Laudanum, celui qu'on loue, raison pour laquelle on le soupçonna d'être lui-même opiomane... Il mourut et sa potion se répandit dans toute l'Europe. D'où est venue l'idée de fumer de l'opium ? Elle est en relation avec l'arrivée du tabac en Chine.

Au début du 19^e siècle, l'allemand Sertürner, isole la morphine, premier alcaloïde chimiquement obtenu. Il l'appela ainsi en raison du Dieu romain du sommeil Morphée. L'utilisation de la morphine se développa avec l'invention de la seringue hypodermique, en 1850, car elle permettait de soulager quasi immédiatement la douleur. Mais le corps médical s'inquiéta rapidement de la forte dépendance que causait la morphine et des études furent entreprises pour éliminer ses propriétés addictives. En effet, elle fut utilisée massivement lors de la guerre de Sécession aux États-Unis entre 1861 et 1865 et pendant la guerre franco-prussienne de 1870 pour le traitement des blessés : les chirurgiens utilisaient la morphine à doses massives pour effectuer des amputations, ignorant tout de son aptitude à créer la dépendance. Des cas de morphinomanie furent d'abord signalés aux États-Unis, en Allemagne, puis un peu partout en France. C'est dans ce contexte que Wright, synthétisa, à Londres, l'héroïne en 1874. En 1898, elle fut commercialisée par la firme pharmaceutique allemande Bayer, pour le traitement de la toux et des douleurs. L'héroïne semblait efficace et mieux encore d'après Dreiser (testeur pour les laboratoires Bayer), elle ne créait pas de dépendance. Il était même question de l'utiliser en produit de substitution de la morphine !!! Bayerregistra ce nouveau médicament sous le nom d'Héroïne de l'allemand, « heroisch », héroïque. Bayer lança une grande campagne marketing et exporta de l'héroïne dans pas moins de 23 pays. En 1911, le British Pharmacoeucical Codex nota que l'héroïne était aussi addictive que la morphine et en 1913, Bayer en arrêta complètement la production. Dès les années 20, un premier texte international réglemente l'usage de l'héroïne. Ce n'est qu'avec la convention unique sur les stupéfiants de 1961 qu'elle sera progressivement interdite dans la plupart des pays.

La French Connection : elle a opéré dans les années 50 et 70 entre la France et les États-Unis, avec Montréal comme plaque tournante. Ce gigantesque trafic d'héroïne regroupait diverses organisations criminelles et un nombre important de trafiquants français, canadiens et américains. À cette époque, l'héroïne, réputée pour être de très grande qualité, était raffinée à Marseille. Le gouvernement américain était impuissant à enrayer ce trafic. Le nouveau président, Richard Nixon, demanda à la France de se donner les moyens d'arrêter les trafiquants et de démanteler les laboratoires clandestins. Il faudra une vingtaine d'années pour éradiquer ce trafic. De cette époque est née l'habitude de s'injecter l'héroïne qui perdure de nos jours alors que l'héroïne qui circule n'est plus sous une forme hydrosoluble. D'après Hostettmann et Pelt ^{77,78}.

L'héroïne base (poudre brune) qui est vendue en Europe est « coupée » essentiellement avec un mélange de paracétamol et de caféine. Il est rapporté que la caféine aide à vaporiser l'héroïne à une plus basse température quand elle est fumée et renforce légèrement l'effet de celle-ci. Le paracétamol est un analgésique dont le point de fusion est similaire à celui de l'héroïne ⁷⁸.

B] La cocaïne

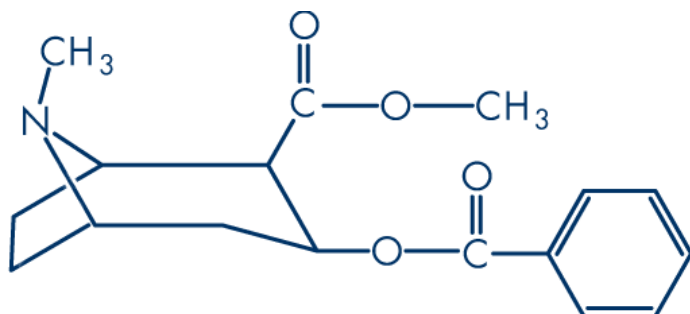
La cocaïne est un alcaloïde extrait de la feuille de coca (*Erythroxylum coca* et *Erythroxylum novogranatense*), qui en contient de 0,5 à 2,5%.

Cet arbuste tropical (figure 43) est cultivé à grande échelle dans la chaîne des Andes en Amérique du Sud et est l'unique source naturelle de cocaïne.



Figure 43 : Cocaïer *Erythroxylum coca* Lam.

Structure moléculaire :



Formule chimique : C₁₇H₂₁NO₄ – poids moléculaire : 303.4 g/mol.

La cocaïne est l'ester méthylique de la benzoylecgonine. Bien que 4 paires d'énantiomères soient théoriquement possibles, un seul énantiomère (généralement appelé l-cocaïne) existe à l'état naturel. La cocaïne est structurellement apparentée à l'atropine (hyoscyamine présente dans la plante belladone) et à l'hyoscine

(scopolamine présente dans la plante datura), des substances possédant des propriétés pharmacologiques très différentes.

L'extraction est réalisée en plusieurs étapes à partir de la feuille de coca séchée : pâte base (pasta), puis cocaïne base et enfin chlorhydrate de cocaïne (figure 44).

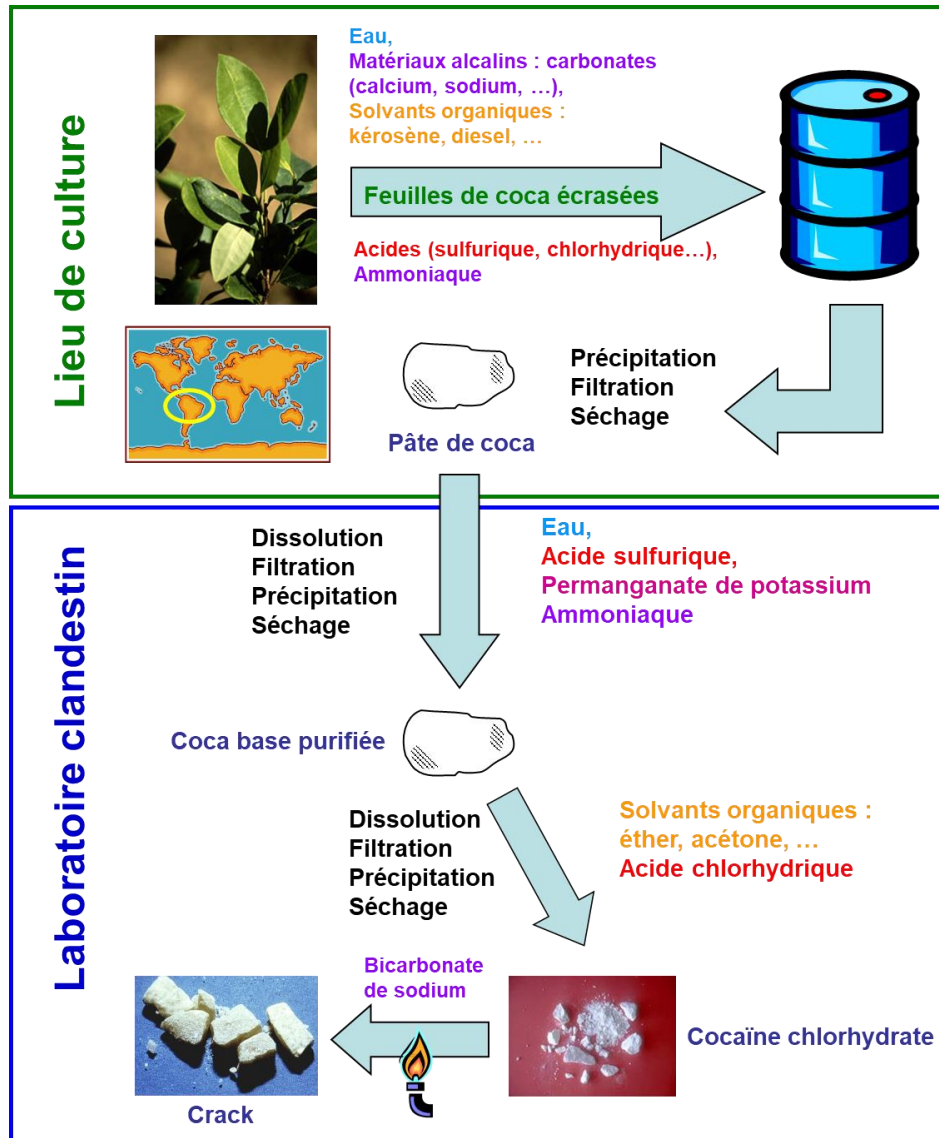


Figure 44 : Étapes de fabrication de la cocaïne à partir de la feuille de coca.

90% de la cocaïne saisie en France est sous forme chimique chlorhydrate. Le crack est de la cocaïne sous forme chimique base libre.

La cocaïne induit un effet psychomoteur stimulant similaire à celui produit par l'amphétamine et les composés apparentés et elle agit également en tant qu'agent anesthésique. Comme l'amphétamine, elle induit de l'euphorie, une tachycardie, une hypertension ainsi qu'une suppression de l'appétit. La cocaïne possède un effet de

renforcement positif important, induisant une rapide dépendance psychologique, effet encore plus marqué chez les personnes qui fument du crack^s.

Petite histoire de la cocaïne

La feuille de cocaïer était mâchée dans les sociétés précolombiennes. Elle est encore présente dans certains pays andins (Bolivie) où la feuille de coca est mastiquée (correspondant à la prise de 75 mg de cocaïne par voie orale) ou bue en infusion pour faire disparaître la sensation de faim et accroître l'endurance et la force de travail. La plante est arrivée en Europe au 18^e siècle. Ce n'est qu'en 1860 qu'a été isolé le principal alcaloïde qu'elle contient par Albert Neumann, chimiste allemand ; l'industrie pharmaceutique l'a lancée peu après sur le marché comme « remède miracle », notamment pour les anesthésies locales. Dès 1860, son usage se répandit en Europe et en Amérique consommée alors sous forme de boisson : en 1863, le chimiste d'origine corse Angelo Mariani commercialisa un extrait de feuilles de coca dans du vin doux sous le nom de « Vin tonique Mariani à la coca du Pérou ». En 1886 est apparu le Coca-Cola®, extrait de feuilles de coca mélangé à de la caféine et de la noix de cola, formule inventée par J.S. Pemberton, pharmacien à Atlanta. Un peu plus tard, la cocaïne a été éliminée de la limonade et a perdu son statut de médicament, tout en gardant celui de drogue en vogue dans certaines minorités (milieu artistique). Mais face au fléau de la cocaïne, les États décidèrent de renforcer les coopérations internationales et adoptèrent la Convention Unique sur les Stupéfiants, signée à New York le 30 mars 1961. La convention place la cocaïne dans la catégorie des stupéfiants et rappelle que toutes les étapes allant de la production à la consommation doivent être contrôlées et limitées à des fins médicales et scientifiques. Dans les années 70 et 80 du siècle dernier, la cocaïne et ses dérivés (freebase, crack) se sont aussi fermement établis sur la scène de la drogue, notamment en association avec l'héroïne (speed-ball, cocktail). Depuis lors, la cocaïne est devenue une drogue à 2 visages: drogue de la misère et drogue de la jet-set. À l'heure actuelle, il n'existe pas de programme de substitution pour la cocaïne. D'après Hostettmann et UNODC ^{77,80}.

Lorsque je travaillais à l'Institut de Police Scientifique et de Criminologie, en Suisse, 2 méthodes de profilage des impuretés majeures de l'héroïne et de la cocaïne étaient déjà en place dont l'objectif était plutôt tactique : c'est-à-dire réaliser des comparaisons entre échantillons de saisie afin de fournir des informations sur d'éventuels liens entre eux pour aider à identifier des réseaux de distribution.

Les méthodes étaient basées sur la séparation de différents composés de l'héroïne et de la cocaïne ainsi qu'une quarantaine d'adulterants et de diluants par chromatographie en phase gazeuse couplée à un détecteur à ionisation de flamme (GC-FID) et/ou à spectrométrie de masse (GC-MS).

Les impuretés majeures détectées pour l'héroïne étaient au nombre de 6 : méconine, acétylcodéine, acétylthébaol, 6-monoacétylmorphine (6-MAM), noscapine et papavérine.

La méconine, la noscapine et la papavérine sont des impuretés alcaloïdes issues de la plante et co-extraites durant l'étape d'extraction.

L'acétylcodéine et l'acétylthébaol) sont des produits d'acétylation issus d'impuretés alcaloïdes et apparaissant au cours de la conversion chimique.

La 6-MAM est un produit de dégradation de l'héroïne et d'impuretés alcaloïdes.

^s https://www.emcdda.europa.eu/publications/drug-profiles/cocaine_fr

Ces impuretés ont l'avantage d'être toujours présentes, seule leur quantité varie d'un lot de production à l'autre.

En ce qui concerne la cocaïne, les impuretés détectées étaient au nombre de 13 : acide benzoïque, anhydroecgonine méthyl ester (AEME), anhydroecgonine (AE), ecgonine méthyl ester (EME), ecgonine, tropacocaïne, benzoylecgonine, norcocaïne, cis-cinnamoylecgonine méthyl ester, acides truxilliques et truxiniques, trans-cinnamoylecgonine méthyl ester et N-formylcocaïne.

L'ecgonine, la tropacocaïne, les truxillines sont des alcaloïdes de la coca. L'acide benzoïque, la benzoylecgonine proviennent de l'hydrolyse de la cocaïne. Les cis et trans cinnamoylecgonine méthyl esters sont des produits d'hydrolyse des cis et trans cinnamoylcocaines. La N-formylcocaïne et la norcocaïne sont des produits d'oxydation de la cocaïne (pas présentes dans la plante). Les AE et AEME sont des artefacts de la chromatographie gazeuse et la EME est trouvée dans la plante, mais n'est pas extraite dans le procédé clandestin, elle est surtout issue de l'hydrolyse de la cocaïne.

Les acides truxilliques et truxiniques proviennent de l'hydrolyse des truxillines ⁷⁹.

L'objectif était alors **d'approfondir l'analyse des données et l'interprétation des liens chimiques**.

Dans cette optique, j'ai obtenu en tant que bénéficiaire principale, un subside de recherche du Fonds National Suisse pour la Science (n° 1115-068017) pendant un 1 an, dont le sujet était « Classification, interpretation and statistical evaluation of data in illicit drug profiling ». Les 2 autres demandeurs étaient les professeurs Franco Taroni et Pierre Margot (directeur de l'Institut à l'époque).

Un rapport final, confidentiel, a été rédigé fin 2003 ⁸⁰, la page de garde et le sommaire sont disponible en annexe 3.

Une première partie a été axée sur l'analyse de chromatogrammes d'héroïne et de cocaïne résultant d'une méthode d'analyse identique, mais obtenus avec 2 instruments différents. L'objectif était de montrer que les liens réalisés sur les 2 appareils étaient les mêmes. Ainsi, en cas de problème sur un appareil, il était possible d'utiliser l'autre. Un autre intérêt de pouvoir utiliser des appareils différents pour établir des liens chimiques est que l'harmonisation d'une méthode d'analyse entre des laboratoires différents, processus compliqué et long cf. projet européen SMT sur le profilage de l'amphétamine, ne serait pas nécessaire.

Sur la figure 45, on constate qu'il y a un décalage des temps de rétention, mais ce n'était pas problématique, car on travaillait sur les aires des pics normalisés par rapport à celle de l'héroïne.

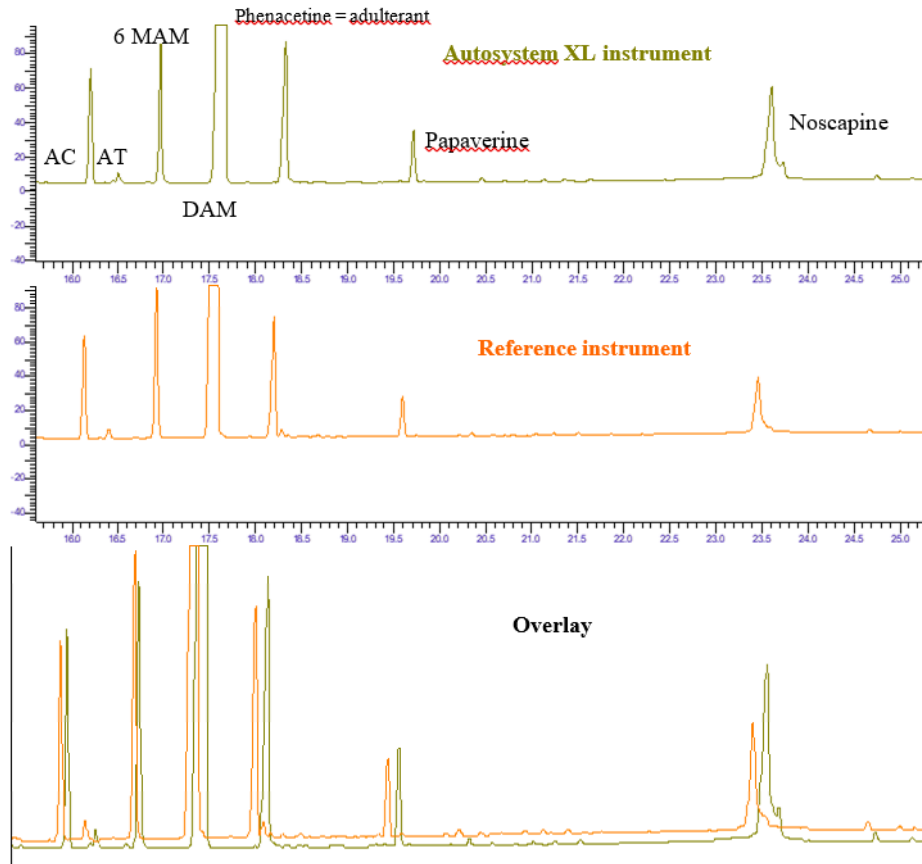


Figure 45 : Chromatogrammes d'un échantillon d'héroïne analysé sur 2 appareils différents.

Le plus problématique était la différence entre les pics de noscapine, sur le deuxième appareil (Autosystem XL instrument), le pic est plus important.

On retrouve ces résultats, figure 46, sur les distributions des aires des pics de la noscapine sur l'ensemble des échantillons analysés.

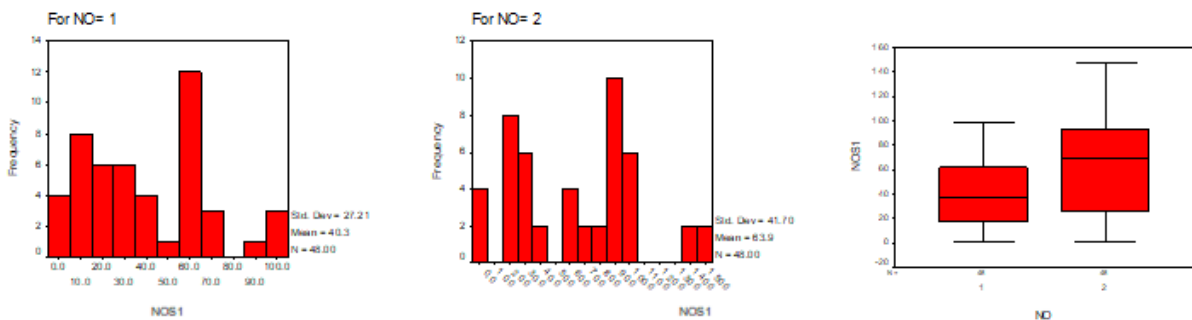


Figure 46 : Distribution et boîtes à moustache des aires de la noscapine pour les 2 appareils.

Pour la cocaïne, on obtenait les mêmes phénomènes notamment au niveau de la benzoyl ecgonine méthyl ester.

L'utilisation de statistiques descriptives permettait de montrer les différences séparément entre des variables. Une approche multivariée pour étudier l'influence des 2 appareils sur l'ensemble des variables dans le cadre de la comparaison de profils chimiques, a ensuite été utilisée. La méthode de comparaison a été décrite dans l'article « A methodology for illicit heroin seizures comparison in a drug intelligence perspective using large databases »⁸¹ dans lequel je suis co-auteur. Le principe est

celui décrit dans la partie sur le profilage en général page 69. La métrique utilisée est la fonction cosinus carré ramenée à une échelle entre 0 et 100, avec 100 la parfaite similitude.

Cette étude a montré qu'il était possible d'effectuer des comparaisons de profils chimiques indépendamment de la machine utilisée pour parvenir à une standardisation "interlaboratoires". En effet, bien que la tâche ne soit pas simple, 2 alternatives sont disponibles. La première concerne l'établissement d'un seuil pour le coefficient de corrélation entre les 2 instruments et c'est la tâche la plus facile, cependant, le risque de faux positifs n'est pas négligeable. La seconde est l'établissement d'une base de données appropriée pour le second instrument, tâche plus longue que la première tâche, mais le problème des faux positifs est évité.

Ainsi, un échange et une évaluation au niveau national des données analytiques sur le profilage de l'héroïne et de la cocaïne pouvaient être réalistes.

Pour confirmer ces résultats et les approfondir, une demande de fonds européen dans le cadre d'une collaboration transfrontalière entre l'institut de Lausanne et nos collègues du laboratoire de Lyon avait été initiée en août 2003, en amont de ce projet.

Le sujet portait sur l'analyse d'échantillons de cocaïne et la comparaison des saisies entre la France et la Suisse afin d'établir des liens chimiques permettant de visualiser le trafic transfrontalier de cocaïne.

Une autre partie de l'étude financée par le FNS concernait le traitement des données de profilage de l'héroïne et de la cocaïne dans un cadre de l'utilisation de base de données importantes. Des méthodes chimiométriques, essentiellement basées sur l'analyse en composantes principales ont été utilisées pour établir des modèles et chercher à savoir s'il était possible d'établir des classes de profils chimiques et ce qui les distinguaient (par ex. une classe d'héroïnes riches en alcaloïdes de la plante, méconine, papavérine et noscapine). L'objectif était d'assigner les profils à des classes dans la base de données et lorsqu'un nouveau profil était obtenu, il était comparé au barycentre des classes et non plus à tous les profils stockés dans la base, ce qui permettait un gain de temps en raison des milliers de profils stockés. Si un « match » était obtenu avec un barycentre d'une classe alors les profils à l'intérieur de cette classe seraient comparés au nouvel entrant pour confirmer son appartenance. La validité de ces classes a été évaluée à l'aide de différents modèles : les réseaux neuronaux artificiels (ANN) et l'analyse discriminante (DA). Pour le modèle ANN, un article intitulé "Chemical profiling and classification of illicit heroin by principal component analysis, squared cosine function and artificial neural networks" a été publié dans Talanta en collaboration avec l'université de Sydney⁸².

La détermination des classes chimiques est particulièrement utile lorsqu'elle est combinée aux informations existantes recueillies par la police. L'enregistrement de ces informations d'une manière significative et utile est crucial pour que les données puissent être utilisées dans une perspective holistique de renseignement.

Mise en évidence de réseaux de distribution

Un réseau de distribution peut être schématisé sommairement selon la figure 47. Comparer, avec des méthodes appropriées, des spécimens au niveau du commerce de détail (de la rue) avec des saisies réalisées à des niveaux plus élevés dans la chaîne permet d'aider à remonter des filières de distribution.

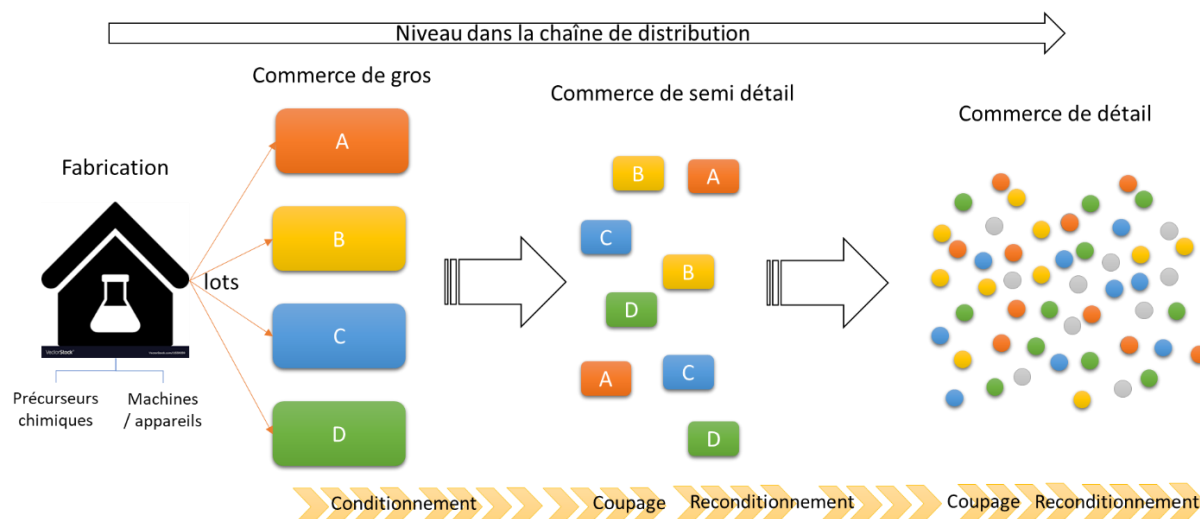


Figure 47 : Mouvement et distribution des drogues illicites du fabricant au consommateur, inspiré de J. Huttunen, et al⁸³.

D'après la figure 47, on peut se retrouver dans différents cas de figure : soit, des liens chimiques sont mis en évidence uniquement entre des saisies de rue (commerce de détail – niveau du consommateur), par exemple les points bleus, on relie alors des consommateurs entre eux et l'information récupérée est qu'ils ont acheté la même drogue et ça s'arrête là. Par contre, si des liens chimiques sont mis en évidence entre des saisies de rue et une saisie ou des saisies au niveau d'un ou plusieurs dealer(s) (commerce de semi-détail), par exemple les petits rectangles bleus, l'information fournie permet d'établir des connexions entre des consommateurs et des dealers. Et lorsque des liens chimiques sont mis en évidence entre des saisies issues du commerce de gros (saisies aéroportuaires ou sur des bateaux), des dealers et des consommateurs, on peut avoir une idée du réseau de distribution bien évidemment en combinant avec des éléments d'enquête.

En conséquence, chaque lien chimique est enregistré dans une base de données d'analyse criminelle de type i2iBase[†]. Cette application permet de combiner les informations issues du profilage chimique avec des informations recueillies au cours des enquêtes et d'afficher ces informations sur un tableau des liens. L'avantage est que le « tableau d'affichage » est le même que celui utilisé par les services de police pour la résolution d'affaires complexes dans lesquelles ils visualisent les liens entre des entités comme des écoutes téléphoniques, des véhicules, des lieux, etc. L'analyse criminelle repose en grande partie sur la représentation visuelle par graphe.

Voici un exemple (figure 48) d'un graphique de liens généré par i2iBase[†] combiné à Analyst Notebook[®] :

[†] <https://i2group.com/>

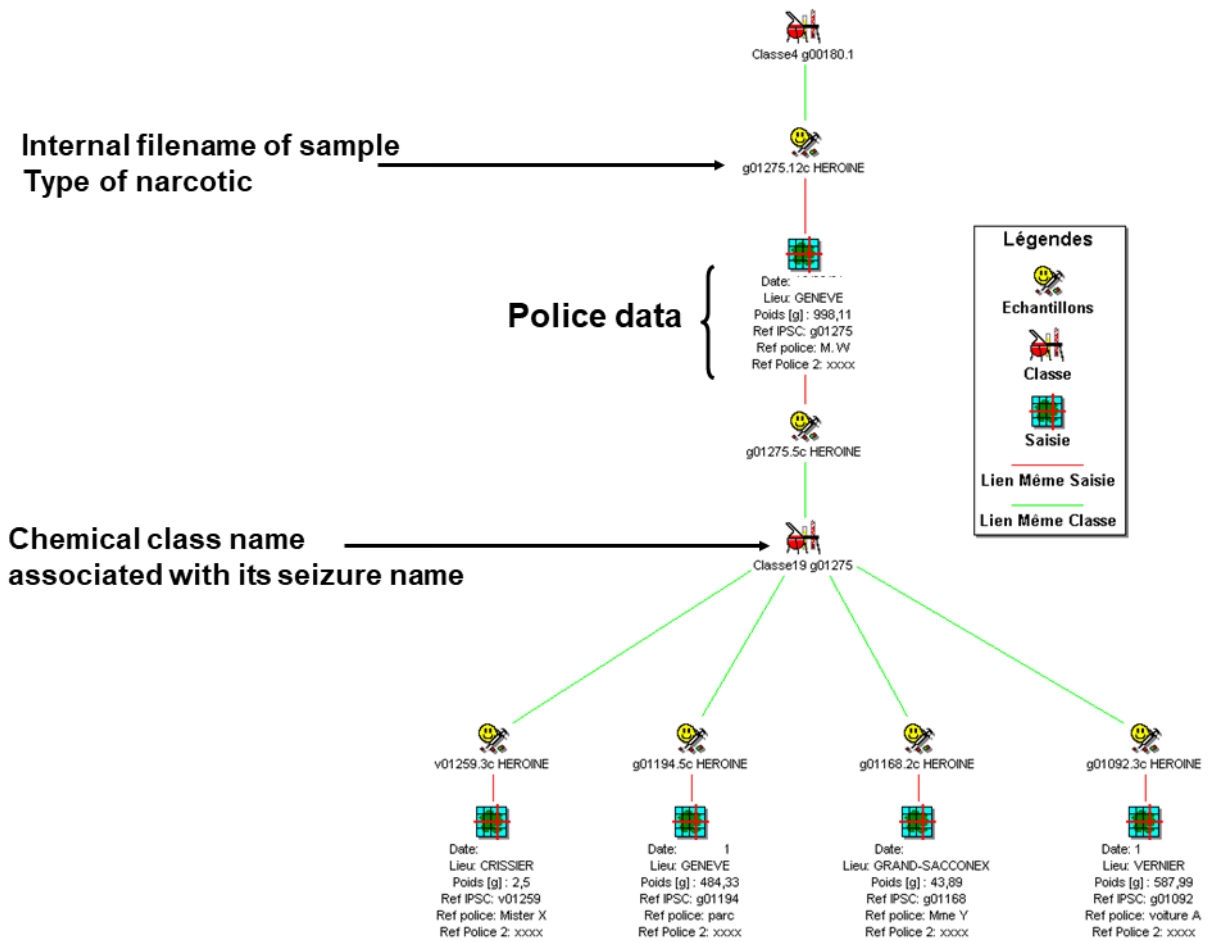


Figure 48 : Principe and visualisation des liens chimiques avec i2 Analyst Notebook® d'après ⁸².

Le projet de collaboration entre l'institut de police scientifique de Lausanne et le laboratoire de police scientifique de Lyon a été financé dans le cadre du programme d'initiative communautaire INTERREG IIIA France/Suisse (2000-2006), n°1/FU/9.3/1 pour un budget total d'environ 310 000€. Il a permis notamment de financer 2 assistants à l'IPS de Lausanne (Mme Stéphanie Lociciro et M. Philippe Hayoz) et l'achat de différentes licences de bases de données.

À cette période, j'avais rejoint le laboratoire de police scientifique de Lyon.

Les **objectifs du projet** étaient multiples :

1. **développement d'un outil de caractérisation et de comparaison** des saisies de cocaïne opérées sur la zone frontalière franco-suisse,
2. **mise en place d'une passerelle informatique sécurisée** via Internet pour la transmission des informations,
3. **observation, classification, interprétation et gestion des liens chimiques** sur les deux régions. Confrontations des liens obtenus avec les renseignements de l'enquête afin d'identifier les réseaux de distribution de la cocaïne,
4. **étude de l'adultération** suivant les deux pays

Les résultats obtenus ont permis de mettre en place une méthodologie de travail pour élaborer une méthode analytique de profilage des saisies de cocaïne commune aux 2 partenaires du projet. Cette étape a nécessité une harmonisation analytique pour comparer les données, rendue délicate à cause des appareillages différents. De

nombreux tests supplémentaires ont donc dû être effectués et une importante étude statistique s'en est suivie afin de proposer une stratégie adéquate de recherche des liens chimiques.

Une passerelle en ligne a été réalisée au moyen d'une base de données FileMaker Pro® et du langage PHP pour la publication sur Internet. Elle a été activement utilisée par les 2 partenaires du projet durant toute la durée du projet pour la transmission des informations et l'exploitation des résultats.

De nombreux liens ont pu être établis grâce aux échantillons analysés en Suisse et en France (environ 20 % des saisies analysées ont présenté un lien Suisse – France et vice versa). Le trafic transfrontalier a donc été démontré par des liens objectifs et validés scientifiquement.

Les éléments d'adultération ont été colligés au sein des 2 laboratoires partenaires. Les produits pouvant être nocifs pour la santé étaient ainsi détectés et l'information transmise rapidement. Ce dispositif d'alerte a d'ailleurs été utilisé avec succès lors de la découverte d'atropine dans des saisies suisses de cocaïne (cristalline).

Plusieurs actions de formation ont été dispensées en France et en Suisse sous forme de cours dans le cadre universitaire de la formation initiale et dans le cadre de la formation continue pour les services de police et les magistrats.

De nombreuses présentations (orales et posters) ont été effectuées en Europe dans le cadre des réunions du réseau européen des instituts de police scientifique (ENFSI) et dans le cadre d'une conférence internationale (IAFS) sur les sciences forensiques à Hong-kong.

La collaboration s'est poursuivie après le projet et a débouché sur **une thèse dont j'ai été co-encadrante** ⁸⁴.

Trois articles ont été publiés ⁸⁵⁻⁸⁷.

Parallèlement, les **recherches ont continué sur le profilage des impuretés de l'héroïne**. En effet, la méthode de profilage utilisée était basée sur les impuretés mineures qui nécessitaient des étapes d'extraction longues et le laboratoire souhaitait basculer sur une méthode plus rapide, discriminante, robuste pour être intégrée à une base de données. Les impuretés majeures de l'héroïne ont servi de base à l'établissement d'une méthode chromatographique couplée directement à une base de données construite sur un serveur web intégrant les calculs mathématiques pour la comparaison. La méthodologie a fait l'objet d'un article intitulé « A quick and automated method for profiling heroin samples for tactical intelligence purposes » publié dans la revue Forensic Science International ⁸⁸.

De plus, une méthode de profilage de la cocaïne à travers l'analyse de solvants résiduels était déjà en place au laboratoire de Lyon et des données étaient enregistrées, mais pas encore totalement exploitées.

Son potentiel pour la discrimination de lots de fabrication et pour la **détermination d'origine géographique a été alors étudié**.

La méthode de détermination des solvants résiduels utilise la chromatographie en phase gazeuse à espace de tête statique (HS-GC) couplée à un détecteur à ionisation de flamme (FID) et est appliquée sur des échantillons de cocaïne chlorhydrate illicite. La préparation de l'échantillon est basée sur un aliquot équivalent à 30 mg de cocaïne pure et ne nécessite donc pas de normalisation par rapport à l'aire du pic de la cocaïne.

Dix-huit solvants résiduels cibles ont été sélectionnés et le résultat des analyses chromatographiques était également transféré directement dans la même base que celle utilisée pour le profilage de l'héroïne.

La figure 49 montre un exemple de chromatogramme obtenu avec la méthode des solvants résiduels. Les pics avec un x suivi d'un numéro (entouré en bleu) sont les solvants cibles.

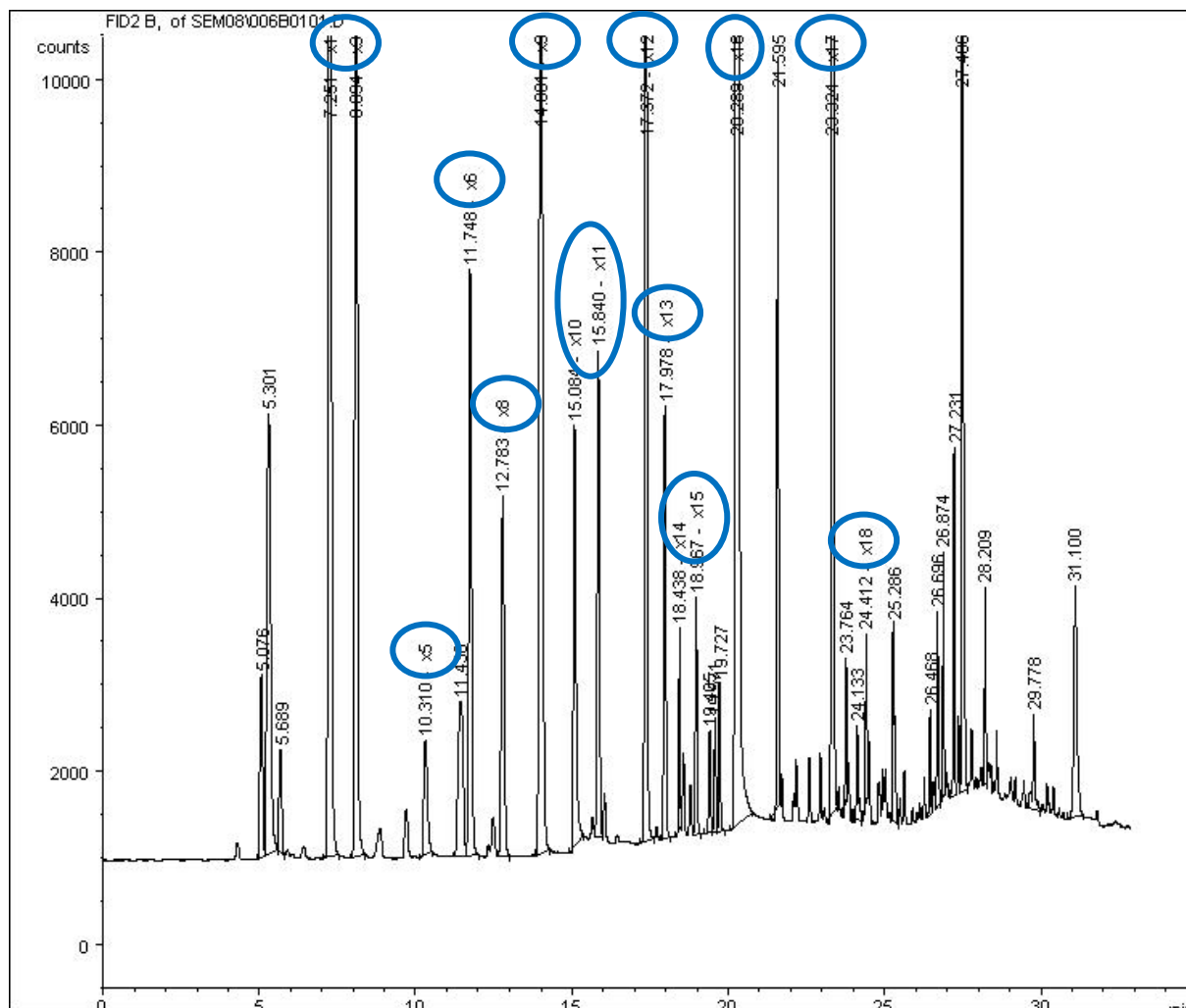


Figure 49 : Exemple d'un chromatogramme d'un échantillon de cocaïne analysé par HS-GC-FID.

Une analyse de classification hiérarchique ascendante (CAH) a été entreprise pour essayer de voir s'il était possible de visualiser les cocaïnes provenant de Bolivie, de Colombie et du Pérou, les 3 principaux fournisseurs. L'analyse a été complétée par une PLS-DA sur ces mêmes échantillons pour connaître les solvants conduisant à la séparation des 3 groupes. La figure 50 montre le graphique des individus et des variables.

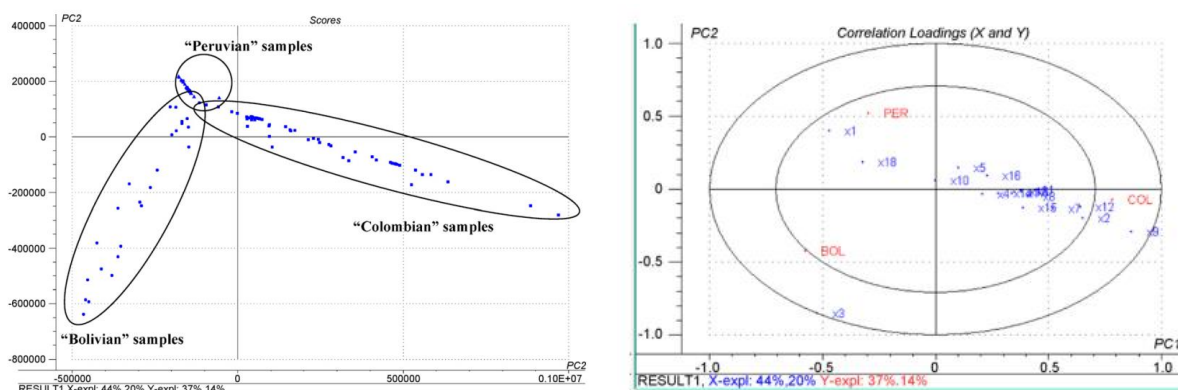


Figure 50 : Graphiques des individus et des variables de la PLS-DA sur les échantillons de cocaïne provenant de Bolivie, Colombie et Pérou.

Les échantillons « boliviens » sont caractérisés essentiellement par de grandes quantités d'acétone (x1), mais aussi de diéthyléther (x3) et du mésityloxyde (x18), mais en faible quantité.

Dans les échantillons "colombiens", ce sont surtout la méthyl-éthyl-cétone MEK (x7), l'hexane ou le chloroforme, ou acétate d'éthyle (x9), le N-propylacétate (x16), et le toluène (x17) très fréquents (>85%) en fortes concentrations qui les caractérisent. L'acétate d'isopropyle ou benzène (x12) est également très fréquent (98%), mais à des concentrations plus faibles.

Dans les échantillons "péruviens", x1 est présent dans 100% des cas en quantité élevée, mais moins que dans les échantillons "boliviens". On note également la présence de x18 (mésityloxyde) et x17 (toluène) respectivement dans 87 et 50% des profils, en faibles quantités. Ces résultats étaient cohérents avec la signature 3 du programme cocaïne de la DEA^u.

Utilisation de l'approche bayésienne pour le profilage des impuretés

La dernière partie de l'étude financée par le FNS a porté sur une réflexion sur la notion de seuil utilisée dans l'évaluation de la mesure d'un lien entre 2 échantillons. Elle s'est d'abord portée sur la comparaison de saisies d'héroïne.

La métrique utilisée pour comparer les échantillons entre eux était la fonction cosinus carré rapportée sur une échelle de 0 à 100, avec 100 correspondant à la similarité parfaite (comparaison de l'échantillon avec lui-même) et 0 la dissimilarité.

Les valeurs de corrélation inférieures à un seuil $S = 99.80$, $C = 99.79$ par exemple, indiquaient que 2 échantillons pouvaient être considérés comme non liés. Et si C était égale à 99.81 par exemple, les 2 échantillons étaient considérés comme liés.

^u Chaque année, dans le cadre du programme de signature de la cocaïne de la DEA (CSP), des analyses chimiques approfondies sont effectuées sur des saisies de cocaïne provenant de saisies en vrac effectuées dans tous les États-Unis. Le programme examine également un plus petit nombre de cocaïne saisies dans le monde entier. En outre, des échantillons de solvants, de réactifs et d'autres matériaux saisis dans des laboratoires illicites de cocaïne en Amérique du Sud sont examinés. Les méthodes d'analyse mises au point au laboratoire de la DEA (STRL) permettent de déterminer comment et où la feuille de coca a été transformée en cocaïne base (origine géographique), et comment la cocaïne base a été transformée en chlorhydrate de cocaïne (méthode de transformation). Les méthodes scientifiques de pointe permettent de déterminer l'origine géographique de la feuille de coca.

L'utilisation d'une approche continue avec un rapport de vraisemblance semblait être une façon d'éviter ce problème de saut de la falaise (fall-off the cliff effect).

Pour cela, une description théorique de la courbe des comparaisons des saisies non liées a été la première préoccupation et l'objectif était d'estimer la densité de probabilités et de connaître le comportement statistique des données. Des distributions Bêta ont été choisies en raison de leur polyvalence et du fait qu'une variété d'incertitudes peut être modélisée avec celles-ci.

La base de données des comparaisons a été coupée en 2 en raison de sa taille importante (> 8000 comparaisons) : la population des valeurs de comparaison comprises entre 22 et 95 et la population entre 95 et 100.

Comme on l'a vu précédemment, on a utilisé un rapport de vraisemblance $LR = P(E|H_1)/P(E|H_2)$.

Dans lequel E est représenté par la médiane de la distribution (un mélange entre les 2 saisies). C'est la médiane de la distribution qui a été choisie, car elle est plus robuste que la moyenne en ce qui concerne les déviations des hypothèses de normalité. H1 et H2 sont représentés par les densités des valeurs de comparaison des saisies λ et γ avec les échantillons stockés dans la base de données non liés (DBNL), respectivement. Sous H1, une saisie est considérée au numérateur (λ ou γ) et sous H2, c'est DBNL qui est au dénominateur.

On calcule ainsi le quotient des valeurs des fonctions de densité de probabilités de la valeur médiane (Med) de la distribution mixte estimée pour les saisies au numérateur (S) et pour les échantillons non liés (DBNL) au dénominateur :

$$LR = \frac{f_{saisie}(Med|H_1)}{f_{DBNL}(Med|H_2)}$$

Équation 6

De grandes valeurs du LR soutiennent le lien et indiquent que les 2 saisies proviennent de la même source. Au fur et à mesure que les valeurs diminuent, la force du lien s'affaiblit jusqu'à des valeurs inférieures à 1, de sorte que le LR soutient que les 2 saisies ne sont pas liées ou qu'elles proviennent de lots chimiques différents. Le graphique de la figure 51 montre un exemple fictif avec une combinaison de saisies non liées en pointillés et liées en gris à droite.

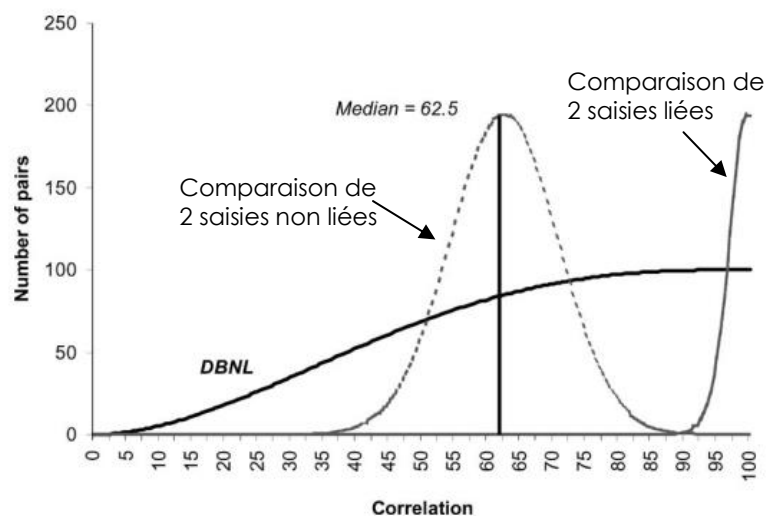


Figure 51 : Exemple fictif du calcul du rapport de vraisemblance avec la comparaison de 2 saisies non liées en pointillés et liées en gris.

Un article détaillant cette démarche a été publié « Evaluation of links in heroin seizures » dans la revue internationale Forensic Science International ⁸⁹.

Dans le cas de la MDMA, 2 approches de rapport de vraisemblance (LR) ont été proposées pour évaluer la force de la preuve des comparaisons de comprimés de MDMA. La première était basée sur une comparaison plus « traditionnelle » des comprimés de MDMA en utilisant des mesures de distance (par exemple, la distance de corrélation de Pearson ou une distance euclidienne). Dans cette approche, les LR étaient calculés en utilisant la distribution des distances entre les comprimés d'un même lot et ceux de lots différents comme dans le cas de l'héroïne décrits ci-devant, mais en utilisant des densités de Kernel gaussiennes (KDE) à la place des distributions Bêta. Ces dernières avaient été aussi testées.

Le rapport de vraisemblance, basé sur 3 caractéristiques physiques, était alors obtenu, non pas par le rapport de la fréquence des caractéristiques elles-mêmes, mais plutôt par la densité de la distance (d) entre 2 mesures x et y lorsque l'une des deux hypothèses concurrentes est vraie :

$$LR = \frac{f(d(x,y)|H_p)}{f(d(x,y)|H_d)} = \frac{f(d(\{x_\emptyset, x_{th}, x_w\}, \{y_\emptyset, y_{th}, y_w\})|H_p)}{f(d(\{x_\emptyset, x_{th}, x_w\}, \{y_\emptyset, y_{th}, y_w\})|H_d)}$$

Équation 7

Avec f la fonction de densité de Kernel gaussienne, \emptyset : diamètre du comprimé en mm, th : épaisseur du comprimé en mm et w : le % massique du comprimé.

H_p : Les comprimés des lots A et B proviennent du même lot.

H_d : Les comprimés des lots A et B proviennent de lots différents.

La deuxième approche était basée sur des méthodes utilisées dans d'autres domaines de la comparaison forensique. Les LR étaient basés sur la distribution des valeurs des caractéristiques des comprimés, plutôt que sur la distribution des distances entre les valeurs des caractéristiques correspondantes de 2 comprimés. Lorsque plus d'une caractéristique est considérée en même temps, on utilise des distributions multivariées au lieu de distributions univariées. Le LR a donc été construit de la même manière que ci-dessus (équation 7), mais sans le terme d.

$$LR = \frac{f(x|y, H_p)}{f(y|H_d)} = \frac{f(\{x_\emptyset, x_{th}, x_w\}|\{y_\emptyset, y_{th}, y_w\}, H_p)}{f(\{y_\emptyset, y_{th}, y_w\}|H_d)}$$

Équation 8

Le calcul du LR a été réalisé en utilisant des distributions normales multivariées pour les mesures au sein d'un lot ($X \sim N(\theta_x, \Sigma_x), Y \sim N((\theta_y, \Sigma_y))$) et une distribution normale multivariée ($\theta_i | \Sigma_i \sim N(\mu_0, T_0)$) pour les moyennes du lot. La normalité au sein d'un lot était une hypothèse raisonnable, mais l'utilisation de KDE pour les moyennes des lots était plus appropriée, ce qui donnait des LR adaptés.

Un article détaillant les 2 approches avec leurs avantages et leurs inconvénients a été publié ⁹⁰.

Exemples de 2 cas concrets analysés au laboratoire de police scientifique de Lyon

a) Profilage des solvants résiduels de la cocaïne

Contexte : en janvier 2006, l'Office central pour la Répression du Trafic Illicite de Stupéfiants repère un cargo battant pavillon panaméen, le Master Endeavour, suspecté de transporter de la cocaïne d'Amérique du Sud jusqu'en Afrique de l'Ouest. Il est arraisonné un mois plus tard le 17 février 2006 par la Marine nationale à 1 300 kilomètres au large des côtes sénégalaises avec plus de 4 tonnes de cocaïne à bord⁹¹. Plus d'une centaine d'échantillons sont envoyés au laboratoire de LPS de Lyon pour analyses (les 3 étapes). En effet, la cocaïne était compactée et emballée sous forme de blocs ou pains de plusieurs kilos et des prélèvements avaient été réalisés par l'OCRIS (Office Central pour la Répression du Trafic Illicite de Stupéfiants devenu l'Office anti-stupéfiants (OFAST) le 1^{er} janvier 2020) sur différents ballots.

À l'issue des étapes 1 et 2, la présence de cocaïne a bien été confirmée (il ne s'agissait pas d'un leurre) et la teneur en principe actif oscillait entre 70 et 85% (teneur très élevée à l'époque et caractéristique d'une cocaïne en provenance directe d'un laboratoire clandestin). Une partie des échantillons ne présentaient aucun produit de coupage et pour d'autres du diltiazem (substance commercialisée pour ses propriétés antihypertensives, anticalcique utilisé en cardiologie) et du lévamisole (antihelminthique (vermifuge)) de la famille des dérivés imidazolés utilisés en médecine vétérinaire) étaient présents en faibles quantités.

À l'issue de la dernière étape d'analyse, des liens chimiques ont pu être mis en évidence :

- D'abord au sein de cette même saisie : les profils de solvants montraient **3 types de lot** qui tout en étant différents présentaient des caractéristiques similaires laissant à penser qu'ils avaient été fabriqués dans le même laboratoire clandestin,
- Et entre des saisies réalisées antérieurement.

La figure 52[A] montre une vue d'ensemble des liens chimiques au sein de la saisie réalisée sur le cargo.

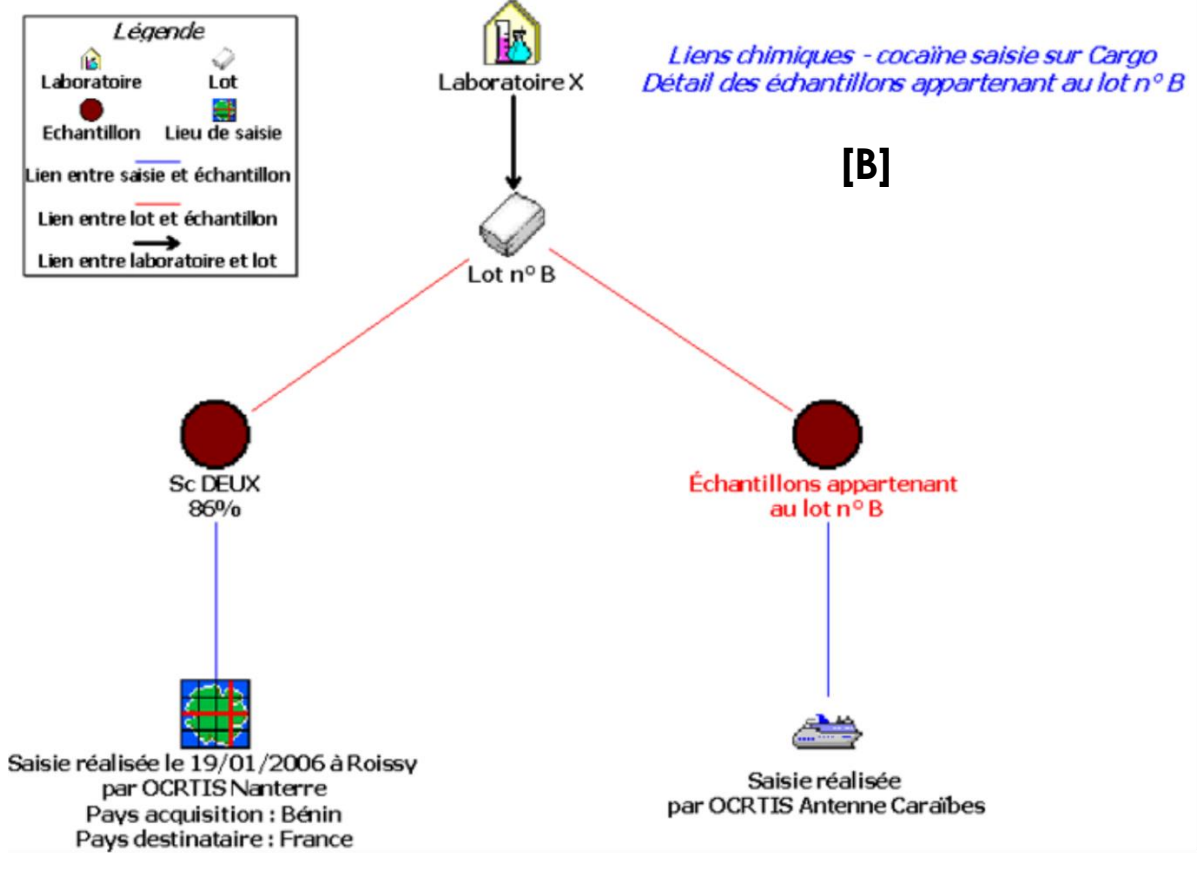
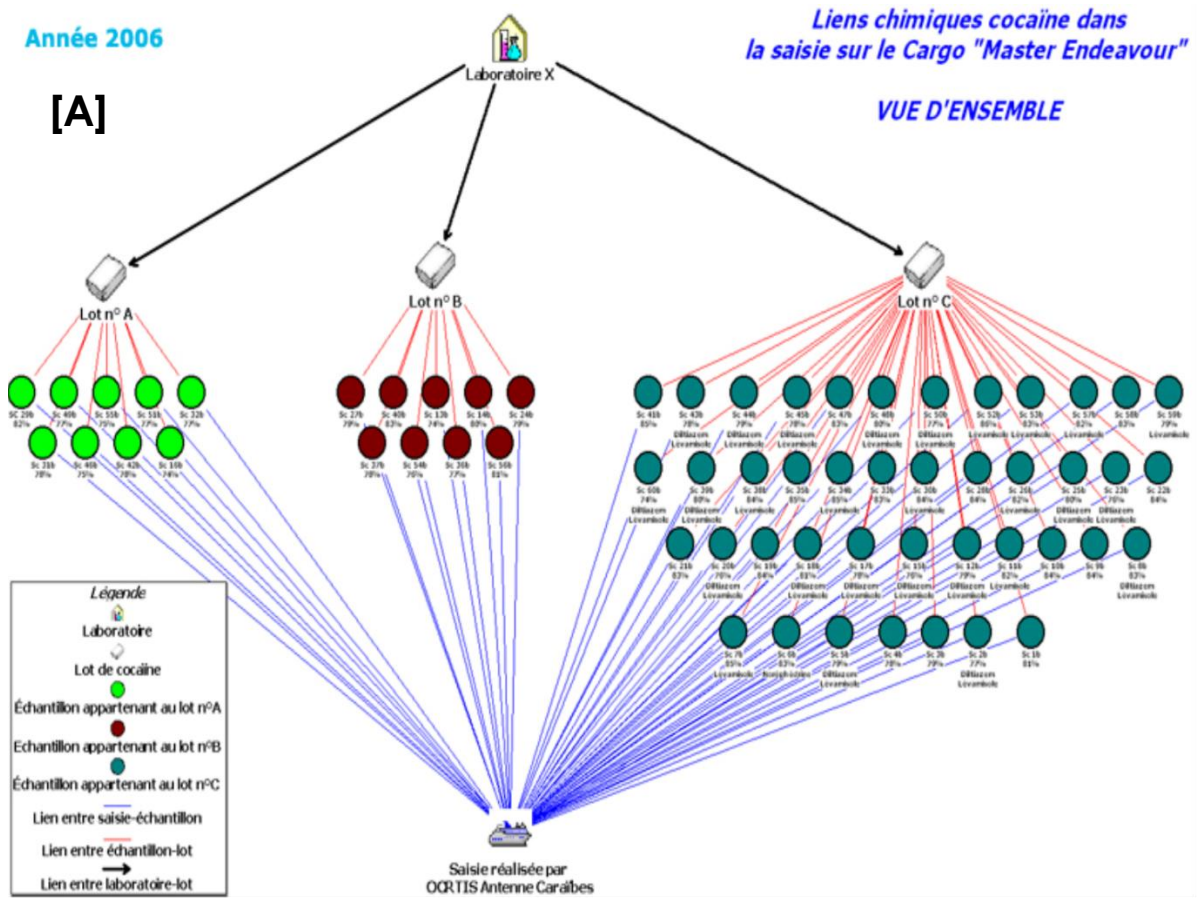
Les lots A et B contiennent chacun 9 échantillons différents et le lot C, le plus important en contient 42. Le lot C est celui dont les échantillons sont adultérés avec du diltiazem et du lévamisole.

Le lot A n'avait pas de lien chimique avec d'autres saisies analysées au laboratoire.

Les figures 52[B] et 52[C] montrent le détail des liens chimiques entre les échantillons des lots B et C, respectivement, et d'autres saisies réalisées antérieurement.

Pour le lot B, un lien chimique a été mis en évidence avec une saisie réalisée à l'aéroport de Roissy en janvier 2006 dont le pays d'acquisition était le Bénin et le pays de destination était la France.

Les échantillons du lot C étaient liés chimiquement à 4 autres saisies, 1 réalisée à l'aéroport de Roissy en janvier 2006 et 3 autres à Nanterre entre janvier et avril 2006. Les pays d'acquisition étaient le Nigéria et le Bénin et les pays de destination l'Espagne et les Pays-Bas.



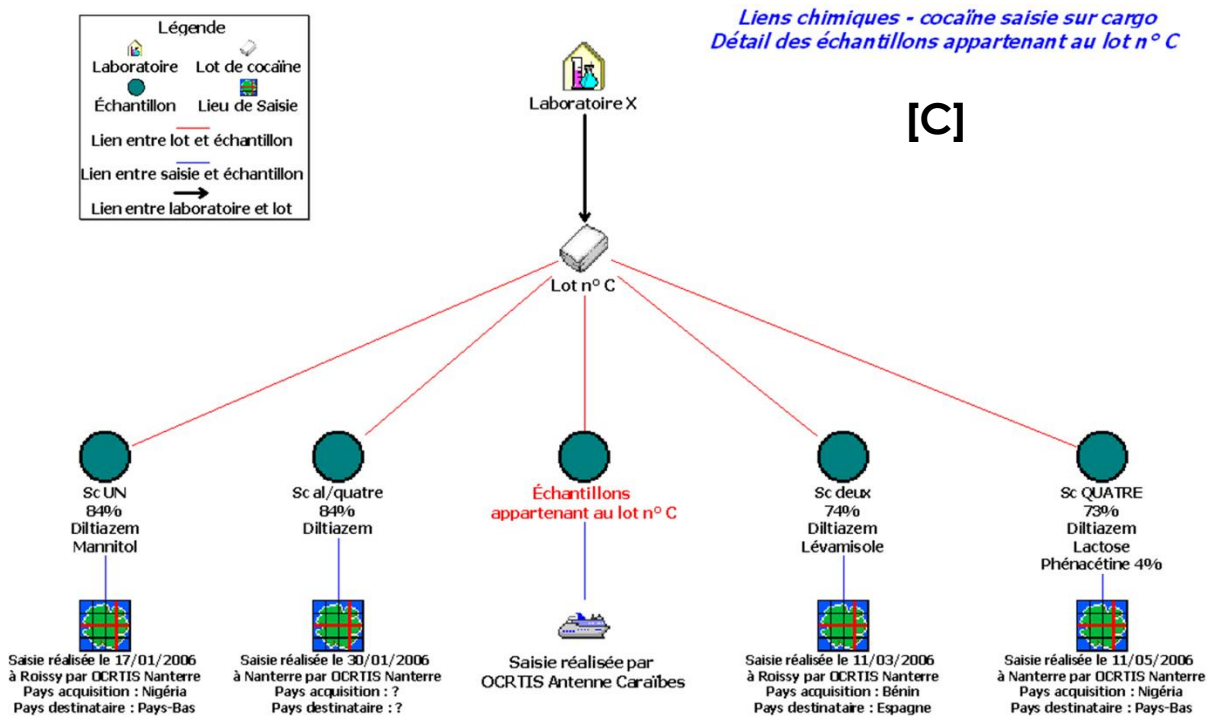


Figure 52 : illustration des liens chimiques mis en évidence dans l'analyse d'échantillons de cocaïne provenant d'une saisie sur un cargo. [A] Vue d'ensemble des 3 lots de fabrication et leurs ramifications. [B] Vue détaillée des liens chimiques pour le lot B. [C] Vue détaillée des liens chimiques pour le lot C.

Cet exemple illustre donc un trafic international de cocaïne en provenance d'Amérique du Sud qui transite par les Caraïbes puis par l'Afrique de l'Ouest pour être enfin distribué en Europe. La figure^v ci-dessous montre les principales routes de la cocaïne. La route pour cette filière est colorée en rouge (figure 53).

^v Issue de <https://www.unodc.org/unodc/en/data-and-analysis/the-cocaine-market.html> consulté le 16/08/2022

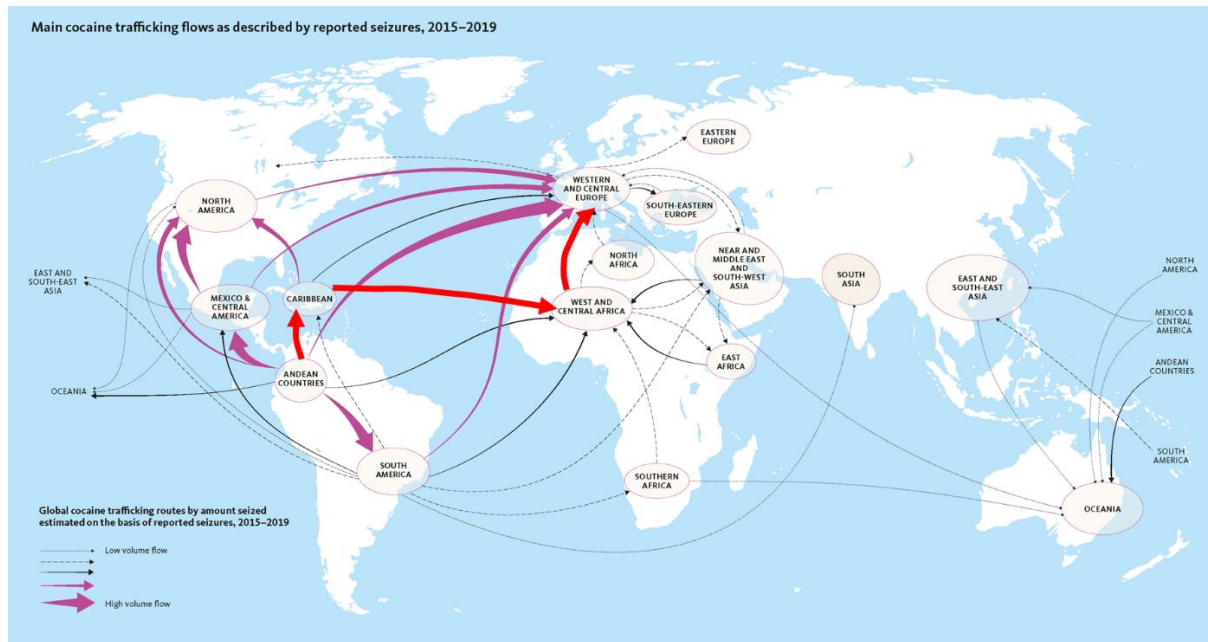


Figure 53 : Principales routes du trafic de la cocaïne, source UNODC.

a) Profilage des impuretés de l'héroïne

Contexte : en début d'année 2007, le LPS de Lyon a été réquisitionné par l'Inspection Générale de la Police Nationale (IGPN) pour réaliser des analyses d'échantillons suspectés d'être de l'héroïne provenant de 2 saisies, l'une composée de 3 échantillons et l'autre de 4. Aucune information supplémentaire n'avait été fournie. À l'issue des analyses, il a été montré qu'il s'agissait bien d'héroïne base (la plus commune circulant en France) dont les teneurs variaient entre 12% (héroïne plutôt de rue) et 50% (héroïne en haut de la chaîne du trafic). Certains échantillons étaient coupés avec du paracétamol et de la caféine (mélange typique rencontré dans l'héroïne base)^w, d'autres non.

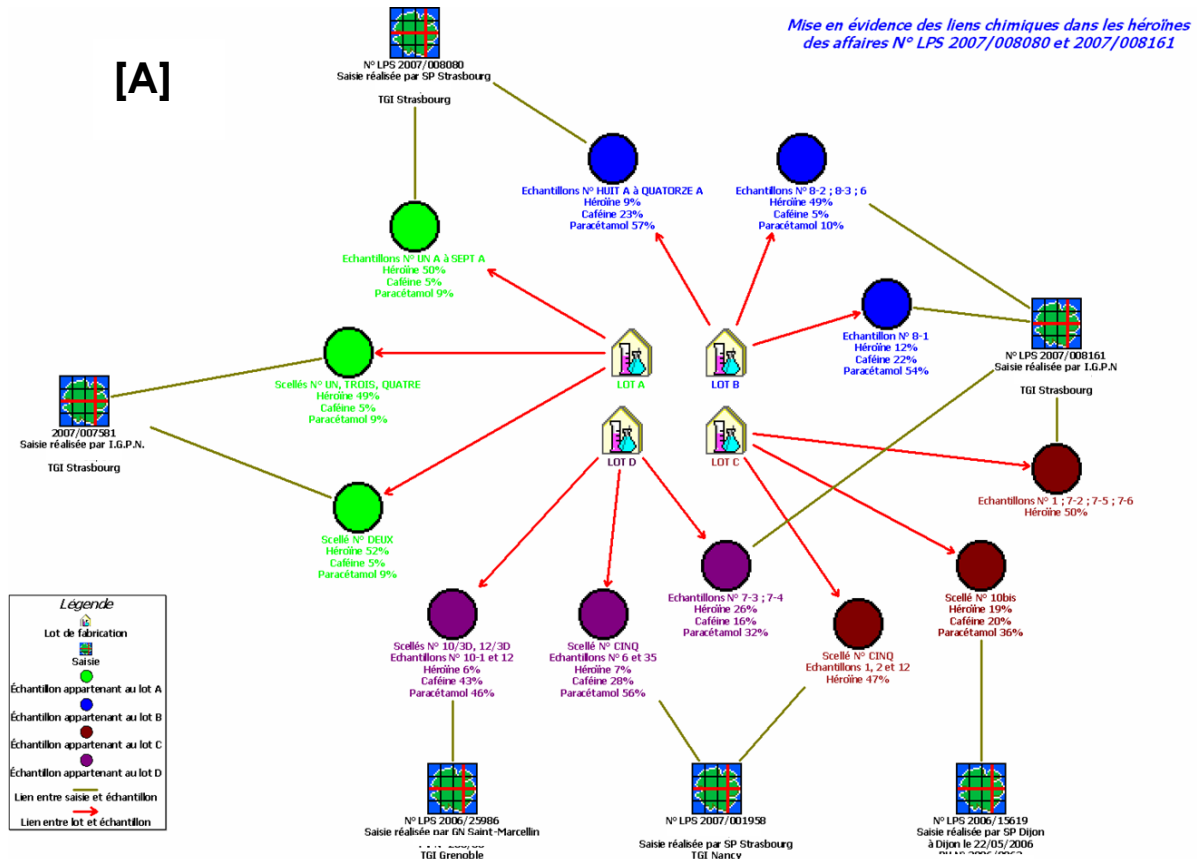
Les analyses de profilage des impuretés de l'héroïne de ces saisies ont révélé **4 lots différents** qui étaient reliés à 4 autres saisies analysées antérieurement.

- Le lot A comprenait les échantillons d'héroïne de la première saisie de l'IGPN et un échantillon d'une saisie réalisée par Sécurité Publique de Strasbourg (en vert clair sur la figure 54A).
- Le lot B comprenait 2 échantillons de la seconde saisie liés chimiquement à un autre échantillon de la saisie réalisée par la SP de Strasbourg (en bleu sur la figure 54A).
- Le lot C était composé d'un échantillon de la 2^{de} saisie de l'IGPN lié chimiquement à 2 échantillons de 2 saisies antérieures, l'une réalisée par la SP de Dijon (saisie à Dijon en mai 2006 et l'autre par la SP de Strasbourg, requête du Tribunal de Grande Instance (TGI) de Nancy (en marron sur la figure 54A).
- Et enfin, le lot D comportait le 4^e échantillon de la 2^{de} saisie de l'IGPN et 2 échantillons de 2 autres saisies, l'un provenant de la SP de Strasbourg, requête

^w Source : <https://www.ofdt.fr/produits-et-addictions/de-z/heroine-et-autres-opiaces> consulté le 16/08/2022

du TGI de Nancy et l'autre, un échantillon saisi par la gendarmerie nationale de Saint-Marcelin, requête du TGI de Grenoble (en violet sur la figure 54[A]).

Après enquête et jugement, il s'est avéré qu'il s'agissait d'un trafic d'héroïne dont le pôle central était Strasbourg et les personnes impliquées étaient une greffière du TGI et un officier de police (cf. encart de presse dans la figure 54[B]).



[B]

Société
Strasbourg: incarcération d'un officier de la brigade des stupéfiants et d'une greffière

Ils ont été mis en examen jeudi soir pour «infraction à la législation sur les stupéfiants». Le premier est poursuivi pour «abus de confiance par officier public» et la seconde, chez qui ont été saisis 4,5 kg d'héroïne, pour «blanchiment».

Par Thomas CALINON (correspondant à Strasbourg)
 LIBERATION.FR : vendredi 30 mars 2007

©IGN Paris 2000

Figure 54 : Flux du trafic d'héroïne dans une affaire en particulier (Strasbourg).

2.7. Conclusion

Ce chapitre rend compte de **16 années de recherche appliquée au domaine de l'analyse de stupéfiants**, de son point de départ, l'échantillon, jusqu'à sa finalité, l'interprétation du résultat dans un contexte légal.

Les analyses se déroulent en 3 étapes dont la première (l'identification) est une étape décisive puisqu'elle conditionne l'établissement d'une infraction au sens du Code pénal et du Code de la santé publique comme le montre la figure 55. Les deux autres étapes (dosage et profilage), aussi importantes et notamment le profilage, peuvent :

- d'une part, apporter un complément d'information à une enquête en cours,
- d'autre part, conditionner l'importance de la peine.

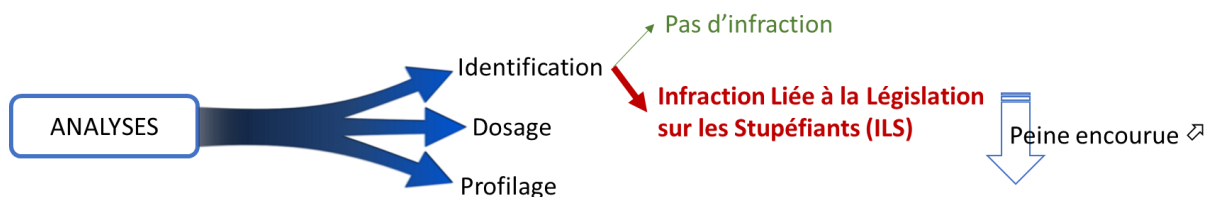


Figure 55 : Trois étapes des analyses dans le cadre supposé d'une infraction ILS et les répercussions que cela peut impliquer pour la personne mise en cause.

Ce domaine est à la croisée de plusieurs champs disciplinaires comme le droit, la géopolitique, la santé, l'histoire des plantes et de leur relation avec l'homme, la chimie analytique et l'analyse de données. Cette diversité le rend « attachant » et cela m'a offert un vaste terrain de jeu dans lequel il reste encore des endroits à défricher et dont les résultats pourraient s'appliquer à d'autres domaines.

Ainsi, l'expérience acquise tout au long de ces années m'a permis de devenir ce qu'on appelle un « expert », c'est-à-dire une personne garante des résultats qui sont fournis par son laboratoire et capable de les défendre devant un tribunal.

Dans le domaine des stupéfiants, il est très rare d'être appelé à la barre en tant qu'expert. Cela se passe à huis clos comme pour les affaires de terrorisme. En général, dans le cas de mise en cause pour du trafic international, les peines peuvent aller jusqu'à la réclusion criminelle à perpétuité. Il m'est arrivé d'être appelée à témoigner à la barre pour une affaire de trafic international de cocaïne au tribunal de Berlin, en 2015, six ans après les faits.

Outre les projets de recherche appliquée mentionnés dans ce chapitre, j'ai également participé à d'autres projets comme des **missions d'audits de laboratoires et de transfert de compétences**, soit dans des laboratoires à l'étranger dans le cadre d'accords bilatéraux entre la France et d'autres pays, soit dans le cadre de jumelages avec des pays demandant l'accession à la Communauté Européenne, ou dans des laboratoires en France (laboratoire du réseau de l'INPS).

J'ai également **participé à la révision et à la rédaction de différents ouvrages** de l'Office des Nations Unies contre la Drogue et le Crime (UNODC) et de l'Observatoire Européen des Drogues et Toxicomanes (EMCDDA), notamment : « Méthodes recommandées pour l'identification et l'analyse des cathinones de synthèse contenues dans des substances saisies », versions française et anglaise 2016⁹², « Methods for Impurity Profiling of heroin and Cocaine »⁴⁵, nouvelle version en cours de publication et « EMCDDA INSIGHTS: Cannabis production and markets in Europe »⁹³.

Durant ce parcours, mon investissement s'est également porté sur **la formation**, d'une part, continue, pour les personnels des laboratoires et les policiers et magistrats

français et étrangers, et d'autre part, initiale, avec des interventions dans divers organismes, dans les domaines de l'interprétation de la preuve scientifique, de l'analyse et du profilage chimiques des drogues ainsi que de l'analyse des données.

Vous devez vous demander comment une experte en analyse de stupéfiants en arrive à travailler dans le domaine de l'agronomie et de l'agroalimentaire ?

Passionnée par la recherche appliquée et l'enseignement, c'est cette envie de partager mes connaissances et de les enrichir qui ont fait que j'ai saisi l'opportunité d'un détachement pour exercer le métier d'Ingénieur de Recherche à l'Institut Agro Dijon avec une composante d'enseignement en statistiques.

Chapitre 3

Réflexions sur les activités de recherche et perspectives

3.1. Changement de cap, des stupéfiants à l'agronomie et à l'agroalimentaire

Les travaux de recherche que j'ai menés durant plusieurs années m'ont permis d'avoir une certaine vision globale du processus d'analyse de substances sous différentes formes physiques (poudre, comprimé, végétal, liquide).

Aujourd'hui, mon travail de recherche est à la jonction entre 2 domaines principaux : la chimie analytique et l'analyse de données. J'essaie d'apporter un nouvel éclairage sur les thématiques traitées à l'Institut Agro Dijon et je continue à explorer différentes façons d'aborder le profilage orienté plus particulièrement vers les méthodes statistiques et chimiométriques.

Arrivée en septembre 2016, il m'a fallu m'adapter au nouvel environnement d'une école d'ingénieurs du ministère de l'Agriculture et de l'Alimentation. Après une période transitoire surtout dédiée à l'enseignement, **j'ai intégré en 2019 la Cellule d'Appui à la Recherche en Science des données au sein de la Direction Scientifique**. C'est une **position transversale** qui a pour objectif d'apporter à la communauté de l'Institut Agro Dijon un savoir-faire en analyse des données dans le cadre de projets de recherche.

Ainsi, j'ai apporté mes compétences en chimiométrie pour différents projets de recherche dans le domaine agroalimentaire, notamment en collaboration avec des chercheurs des **UMR Procédés Alimentaires et Microbiologiques (PAM) et Centre des Sciences du Goût et de l'Alimentation (CSGA)**.

Avec l'**UMR PAM** et l'équipe de la **Physico-chimie des Aliments et du Vin (PCAV)**, j'ai travaillé sur des données issues principalement de chimie analytique comme la chromatographie gazeuse couplée à la spectrométrie de masse, la spectroscopie infrarouge, micro-ondes et sur des thématiques variées comme :

- La détection de fongicides dans les vins à l'aide de polymères à empreinte moléculaire (MIS) couplés à un capteur micro-ondes. L'objectif était d'explorer le potentiel de méthodes chimiométriques pour interpréter les spectres micro-ondes et pour corréliser les signaux micro-ondes avec les concentrations de fongicides d'abord dans un milieu de vin modèle puis dans des vins de Bourgogne blancs et rouges ;
- L'analyse rapide et directe des niveaux d'oxydation des lipides dans des aliments et boissons. Il s'agissait de déterminer une méthode efficace et simple pour surveiller l'oxydation des acides gras oméga-3 dans des émulsions d'huile de poisson dans l'eau (20% v/v). La spectroscopie infrarouge à transformée de Fourier (FTIR) a été utilisée pour enregistrer directement les changements chimiques survenant au cours de l'oxydation des lipides. Cette dernière a été suivie et quantifiée en utilisant les valeurs de diènes conjugués (CD) après une extraction des lipides. Des modèles de régression par les moindres carrés partiels (PLSR) ont été établis en utilisant les spectres entiers (4000-700 cm⁻¹), la région carbonyle (1800-1500 cm⁻¹) et les régions « fingerprint » (1500-700 cm⁻¹).

Ces modèles ont fourni une bonne prédiction des valeurs de CD avec une erreur moyenne (moyenne de RMSECV) de moins de 18% ;

- L'exposition de la population libanaise à des éléments en traces (TRs : As, Cd, Co, Cr, Hg, Ni, et Pb) dans le pain pita. Cette étude était orientée, d'une part, sur l'exploitation de données d'une vaste enquête sur un échantillon de 1000 répondants représentatif des consommateurs de pain pita au Liban et, d'autre part, sur l'analyse des prélèvements des 3 marques de pita (B1, B2 et B3) les plus consommées, par spectroscopie d'absorption atomique (SAA). Une caractérisation du risque sanitaire a été proposée ;
- La validation d'une nouvelle méthode de suivi des éléments traces dans les sols céréaliers méditerranéens. La méthode de référence utilisée pour surveiller les quantités d'éléments en traces dans le sol fait appel à des acides puissants (acide perchlorique (HClO₄) et à l'acide fluorhydrique (HF) et à un appareillage coûteux (spectrométrie d'émission optique à plasma à couplage inductif (ICP-OES) et de spectrométrie de masse à plasma à couplage inductif (ICP-MS), ce qui ne permet pas à certains laboratoires de réaliser ce type d'analyse. Une méthode alternative basée sur la digestion sulfurique par micro-ondes (MSD) combinée à la spectroscopie d'absorption atomique (AAS) a donc été proposée et il s'agissait donc de valider cette nouvelle méthode quantitative et de montrer qu'elle pouvait être utilisée en routine ;
- Une étude de la régionalité des vins australiens issus du cépage Pinot noir par Résonance Magnétique Nucléaire (RMN) et par spectrométrie de masse à plasma à couplage inductif (ICP-MS). L'idée était d'utiliser les données spectrales des 2 techniques pour les combiner dans des analyses multitableaux (analyse factorielle multiple MFA et de coinertie MCOA). Il est apparu que l'ICP-MS semblait être la technique de choix pour déterminer le caractère régional des vins australiens et néo-zélandais, reflétant peut-être les extrêmes de la géologie que l'on trouve souvent dans ces 2 nations.
- L'identification des composés volatils dans les baies de cassis et la possibilité de les discriminer selon leurs variétés. Une première analyse de métadonnées issues de la littérature a été réalisée afin d'étudier les paramètres pertinents pour différencier les baies selon leurs composés aromatiques et la variété à laquelle ils appartiennent. Ensuite, les fractions volatiles ont été extraites et analysées par HS-SPME-GC-MS sur des baies récoltées sur des parcelles appartenant à la chambre d'agriculture de la région Bourgogne. La démarche utilisée pour le profilage chimique des drogues a été adaptée dans ce cas précis et a permis d'établir des groupes au sein des différentes variétés.

La valorisation de ces différents projets s'est soldée par 7 articles ⁹⁴⁻¹⁰⁰ dans le cadre de 2 thèses ^{101,102}.

J'ai également collaboré avec l'équipe **Procédé Microbiologiques et Biotechnologiques (PMB)** sur des données provenant d'analyses spectroscopiques infrarouges sur des levures et des spores. Les objectifs étaient de comprendre les effets du stress sur ces formes de vie. En effet, les spores bactériennes sont très résistantes et jouent un rôle important dans la détérioration des aliments et les maladies d'origine

alimentaire. En ce qui concerne les levures, les cellules sont souvent soumises, pendant leur production industrielle, à des variations hydriques délétères lors de la déshydratation, réduisant ainsi leur viabilité et leur activité cellulaire.

Des expériences mettant en œuvre différentes façons de générer du stress sur les cellules (traitement à haute pression pour les spores, ou déshydratation pour les levures) ont été réalisées et des analyses en microspectroscopie infrarouge à transformée de Fourier (FTIR) avec rayonnement synchrotron (S-FTIR) ont été entreprises pour étudier les réponses biophysiques (mécanismes de résistance) de ces cellules à un stress et essayer de déterminer des signatures biochimiques. L'utilisation de la chimiométrie comme outil exploratoire a permis de cerner les zones spectrales des lipides et des amides, dans le cas des spores conduisant à la discrimination de plusieurs types de cellules (par exemple, des spores germées, mortes, dormantes), traitées à haute pression. Dans le cas des levures, les zones caractéristiques étaient celles des lipides et des protéines et des groupes distincts ont pu être mis en évidence pour des cinétiques de déshydratation différentes.

Les résultats obtenus dans le cadre de **2 thèses ont été valorisés sous forme de 2 articles** ^{103,104} **et un poster** à IAFP'S European Symposium on Food Safety (cf. N° 16 dans la partie du CV).

Avec le **Centre des Sciences du Goût et de l'Alimentation (CSGA)**, j'ai collaboré aussi dans le cadre **d'une thèse** ¹⁰⁵ poursuivie par un postdoctorat sur une étude exploratoire combinant des données de réalité virtuelle et d'oculométrie (eye-tracking) sur le comportement du regard envers les légumineuses dans les rayons d'un supermarché virtuel. Le travail s'est découpé en plusieurs parties, la première a consisté à construire l'environnement virtuel sur différentes étagères : les produits de 4 rayons différents devaient être répartis de façon équilibrée en fonction du groupe alimentaire auquel ils appartenaient c.-à-d. les rayons "conserves", "secs" et "réfrigérés" contenaient des produits d'un seul groupe alimentaire (féculents, produits d'origine animale, légumes secs ou légumes), non transformés ou transformés. La seconde partie concernait l'organisation de l'expérience virtuelle avec la récolte des données et enfin l'exploitation de celles-ci. Il a été montré que les mesures implicites par « eye-tracking » permettaient d'identifier le comportement du regard des consommateurs français envers les légumes secs dans les différents rayons du supermarché virtuel, sans avoir à les interroger directement. Il a été constaté que les légumes secs n'attiraient pas le regard, en particulier dans les rayons des produits secs et des conserves. De plus, les résultats ont également montré qu'il existait une certaine relation entre le comportement du regard et les choix, mais que ce lien était plus complexe que prévu. Dans l'étude, non seulement la durée de fixation, mais aussi les motivations (selon le scénario présenté), et le groupe alimentaire auquel appartient le produit ont influencé les choix alimentaires des participants. Trois tendances différentes pour la relation entre le regard et le choix, en fonction de la motivation, ont pu être observées : (i) une faible fixation sur un groupe de produits, mais un choix très fréquent de ces produits ; (ii) des fixations fréquentes, mais un choix peu fréquent d'un groupe de produits ; (iii) aucune relation entre les fixations et le choix, où des fréquences de fixation similaires conduisent à des fréquences de choix différentes.

La **valorisation des résultats a conduit à 2 articles** ^{106,107} **et 3 participations à des congrès internationaux** : au 14^e Pangborn Sensory Science Symposium (cf. n° 12 dans

la partie du CV) et à la 9^e European conference on sensory and consumer research (Eurosense) (cf. n° 13 et 14 dans la partie du CV).

De plus, un article est en cours de rédaction sur l'étude des intentions d'acheter et la volonté de goûter de nouveaux produits à base de légumineuses des consommateurs.

3.2 Perspectives à court et moyen terme

Les travaux de recherche que je souhaite réaliser pour la suite s'orientent sur **l'utilisation astucieuse et combinée de méthodes statistiques/mathématiques existantes**. Il existe, en effet, des méthodes efficaces pour traiter et extraire des informations pertinentes à partir de données d'analyses chromatographiques, spectrales, etc., pour détecter des pics, identifier des composés cibles, repérer des tendances, etc. Une partie de ce mémoire en fait état.

Mais il y a aussi des images avec l'extraction de caractéristiques, ou encore des données textuelles, des données mixtes (qualitatives et quantitatives) qui, de plus en plus, représentent de gros volumes.

Chacune de ces sphères fournira son lot de questions intéressantes et la nécessité de développer des méthodologies visant à améliorer l'existant. En effet, **combinaison des données, combinaison des résultats des analyses statistiques de chacun de ces types de données dans un objectif d'interprétation dans un contexte donné** est une perspective que je souhaite suivre dans mes futures activités.

3.2.1 À court terme

En 2021, un fonds de soutien a été proposé par l'Institut Agro Dijon (anciennement AgroSup Dijon) pour renforcer sa politique de recherche. Il avait vocation à financer des projets portant sur les orientations scientifiques détaillées dans sa stratégie de recherche et d'innovation. L'incitation spécifique de cet appel à projets portait sur la mise en œuvre de projets transversaux répondant aux orientations scientifiques de l'Institut Agro Dijon et des projets d'amorçage à caractère exploratoire.

Dans le cadre de l'axe stratégique en Science des données - « Transition numérique au service des systèmes agroalimentaires durables », un projet transversal dans le lequel je participe en tant que chercheuse associée a été retenu. Il s'intitule « **IAforStressLife** » - Introduction de l'intelligence artificielle dans la recherche « procédés ». Les structures impliquées sont : l'UMR PAM, équipe PMB, la plateforme DImaCell, le Laboratoire d'Informatique de Bourgogne (IMB) et la Direction Scientifique, Cellule d'appui à la recherche en science des données. Le porteur est le professeur Pierre-Yves Louis, les co-porteurs sont les professeurs Jean-Marie Perrier-Cornet et Pierre-André Maréchal (UMR PAM, équipe PMB).

Ce projet vise à mettre en place une ligne d'acquisition d'images microscopiques de cellules d'intérêt vivantes (type levures par exemple) auxquelles une perturbation contrôlée sera appliquée. Le contrôle et la régulation de cette ligne seront gérés en temps réel par un système choisi d'intelligence artificielle préalablement entraîné sur des données mesurées et/ou sur des données simulées via un modèle mathématique basé sur la physique. Cette chaîne de traitement devrait permettre de relier les caractéristiques morphologiques des cellules transmises par le microscope pendant le traitement avec leurs fonctionnalités après traitement afin de pouvoir adapter en temps réel la perturbation et ainsi de réduire les impacts de cette perturbation.

La variable d'action choisie initialement est la déshydratation osmotique (soluté osmodépresseur), car l'équipe PMB possède une grande expertise dans le domaine

et qu'elle garde un intérêt fort dans la possibilité de préserver des cellules vivantes à température ambiante. Cela permet dans un premier temps de se focaliser sur les développements de la chaîne d'analyse et d'optimisation et moins sur la partie biologique qui est déjà bien maîtrisée par le laboratoire. Cependant d'autres variables d'entrée pourront facilement être utilisées sur la même installation sans modifications importantes : injection d'un soluté actif (drogue, antimicrobien ciblé ...). Ce projet a démarré début décembre 2021 et se déroulera sur 36 mois. Il s'appuie sur une **thèse de doctorat**.

La thèse de M. Frank Hyvrier porte sur la **modélisation augmentée par l'intelligence artificielle pour comprendre et contrôler le processus de déshydratation des micro-organismes de type levures**. Elle s'insère dans les 5 « work packages » ci-après et plus particulièrement dans les WP3, WP4 et WP5 :

- WP1 : mise en place d'une chaîne automatisée d'action/acquisition/régulation de perturbation cellulaire sous microscope,
- WP2 : mise au point d'un système d'analyse d'images permettant de remonter un ensemble de critères pertinents,
- WP3 : développement d'un modèle mathématique, basé physique, ajustable à des données expérimentales,
- WP4 : contrôle du modèle, intégration de la boucle de rétroaction,
- WP5 : développement d'un système de réseau de neurones afin de relier la variable d'action et la viabilité.

L'objectif est d'apporter à l'Institut Agro Dijon (IAD) une compétence dans l'intégration d'intelligence artificielle dans la recherche, domaine très prometteur pour quasiment tous les domaines de recherche de l'IAD. Ce projet original pourrait intégrer, à court terme, d'autres chercheurs et déboucher sur d'autres projets d'ambitions nationales et internationales, dans différents domaines : procédés alimentaires, santé, biologie végétale...

Toujours dans l'Appel à projets de l'Institut Agro Dijon (IAD), **j'anime et coordonne un projet d'amorçage à caractère exploratoire** intitulé « Analyses croisées des pratiques culturelles et œnologiques et des problèmes de fermentation alcoolique sur vins rouges de Bourgogne ». Il s'agit d'un projet transversal ancré dans la filière vigne et vin qui rassemble 6 équipes de différentes disciplines (mathématiques et modélisation/données et sciences humaines et sociales). Il est d'une durée de 24 mois et s'inscrit dans la stratégie de la Recherche et de l'Innovation d'IAD sur 2 axes en particulier :

- L'axe 4 : Former et informer pour des systèmes agroalimentaires durables et plus particulièrement « Accompagner les professionnels (agriculteurs, conseillers, transformateurs) »,
- L'axe 5 : Science des données - applications à la transition vers des systèmes agroalimentaires territorialisés et durables et notamment « Analyser des données en interdépendance et Développer de l'information construite et pertinente pour aider à la décision ».

La fermentation alcoolique constitue une étape clé du processus de vinification et sa maîtrise reste sensible et soumise à l'action de nombreux facteurs, à la fois culturels et œnologiques, influençant positivement ou négativement son déroulement. Il a été observé que des difficultés de fermentation alcoolique s'étaient progressivement étendues au point qu'en 2018 puis 2020 notamment, plusieurs viticulteurs se sont

retrouvés confrontés à des fermentations alcooliques sur vins rouges très ralenties, voire interrompues, configuration en rupture avec les incidents ayant pu survenir dans le passé de par leur fréquence et leur caractère aléatoire, cela quels que soient les itinéraires techniques, à la parcelle ou à la cuverie, adoptés par les intéressés.

Les objectifs scientifiques sont divisés en 3 « work packages » : le premier s'attache à étudier les facteurs liés aux techniques viticoles œnologiques par une 1^{re} approche par questionnaire et une exploitation quantitative des réponses. Les analyses usuelles de traitement d'enquête via des méthodes factorielles de type ACM ¹⁰⁸ et de clustering ¹⁰⁹ afin d'établir des typologies (agrégation d'observations en catégories homogènes) de pratiques montrant des problèmes de fermentation alcoolique (ou non) ^{96,110}.

Des transformations des données originales par recodification ou calcul sont prévues pour affiner ces typologies. Bien plus, une réflexion sera engagée sur la création d'un ou plusieurs indices (de fermentescibilité par exemple). Par ailleurs, des méthodes statistiques, d'imputation de données manquantes seront probablement nécessaires ¹¹¹.

Le second WP s'intéresse aux facteurs liés aux pratiques professionnelles des viticulteurs par une approche qualitative par entretiens et observations. Des traitements de type « text mining » sont envisagés notamment en utilisant des représentations sous forme de graphes pondérés : les mots forment les nœuds et une arête entre 2 nœuds est pondérée par le nombre d'entretiens qui les utilisent conjointement ^{112,113}. Ces graphes permettront une visualisation aisée pour les non spécialistes ce qui facilitera le dialogue. Ces graphes de mots permettront d'interroger les pratiques et de faire ressortir des thèmes ¹¹⁴. Le 3^e WP porte sur l'intégration des résultats des analyses qualitatives et quantitatives. De plus, les informations récoltées, que ce soit par l'intermédiaire du questionnaire ou des entretiens, pourront être croisées avec d'autres variables comme la localisation géographique par exemple. À nouveau, des modèles statistiques sur réseaux pourront être utilisés afin d'explorer les données et dégager une compréhension du phénomène.

3.2. A moyen terme

Le label Carnot Qualiment finance des projets de recherche menés au sein de ses unités de recherche pour anticiper et préparer les besoins de recherche et d'innovation des entreprises agroalimentaires.

Dans le cadre de son édition 2022, **un projet d'une durée de 24 mois**, associant 3 partenaires dont l'IAD et dans lequel je suis associée en tant que chercheuse a été proposé, il s'intitule « DATAVEG ».

L'étude vise à obtenir des données technofonctionnelles sur des protéines végétales. En effet, c'est une donnée difficile à obtenir de la part des fournisseurs, car elle est dépendante de la matrice dans laquelle elle est destinée à être incorporée. Ainsi, sur une fiche technique d'une protéine végétale ne figurent souvent que des critères de composition.

Ce projet consiste donc à travailler sur 3 modèles de matrices alimentaires (1/ Boissons & dessert végétaux, 2/ Analogues de viandes et 3/ Produits céréaliers) et à évaluer la possibilité de mettre en place, pour chacune de ces matrices, des tests applicatifs standardisés permettant de quantifier l'aptitude technologique ; tout en collectant l'ensemble des données d'entrée déjà disponibles (composition physico-chimique, source végétale, procédé d'obtention...). L'objectif final est de construire les premières briques d'une base de données répertoriant l'offre globale en protéines

végétales avec pour chacune d'elle, son score d'aptitude à répondre aux contraintes technologiques de la matrice dans laquelle il est envisagé de l'incorporer.

Mon apport se situe au niveau de l'établissement d'un cahier des charges pour la **construction d'une base de données techno fonctionnelle** des protéines avec l'établissement d'un outil d'aide à la décision intégré à la BDD. Cet outil d'aide à la décision sera basé sur des **méthodes chimiométriques avancées** avec des modèles multiblocs. Ils sont considérés comme des outils d'exploration de données dans la mesure où ils peuvent donner un aperçu (graphique) de grandes quantités de données, dans le but d'améliorer les connaissances sur le sujet étudié en réduisant notamment la complexité. Il existe différentes familles de méthodes suivant l'objectif poursuivi : classification, prédiction, recherche de redondances, sélection de variables, etc.).

Je travaille également avec des chercheurs du **Centre des Sciences du Goût et de l'Alimentation** (CSGA), Stéphanie Chambaron (DR INRAE) et Marie Claude Brindisi (PU, PH, CSGA, CHU), sur des questions autour de l'obésité notamment sur le rôle des fonctions exécutives et de la charge cognitive afin de mieux comprendre leurs rôles respectifs dans le comportement alimentaire des adultes en surpoids et sur l'utilisation du niveau de pleine conscience comme indicateur de l'adhésion aux stratégies de prévention de la prise de poids basées sur l'entraînement à l'autocontrôle avec le neurofeedback. **Des propositions de projet** dans le cadre d'appels de type ANR sont régulièrement envoyées.

En outre, je serai impliquée dans une thèse co-dirigée par Gaëlle Arvisenet (PR Uni) et Stéphanie Chambaron (DR INRAE) à partir de début octobre sur le sujet : « Vers des choix alimentaires plus durables : que comprennent les consommateurs ? » Compréhension des mécanismes perceptifs et cognitifs sous-tendant des choix alimentaires durables.

La question de recherche principale de ce projet consiste à étudier et à comprendre les mécanismes perceptifs et cognitifs sous-tendant des choix alimentaires durables. Le point de départ indispensable à toute stratégie de changement comportemental est d'arriver à comprendre les raisons qui font que les individus adoptent des comportements non optimaux, aux conséquences pas toujours en phase avec leurs préoccupations déclarées.

Ce travail de thèse aura trois objectifs principaux :

- 1) Exploration des représentations mentales des consommateurs concernant la durabilité : identifier les connaissances des consommateurs sur la notion d'alimentation durable, par des méthodologies directes et indirectes.
- 2) Impact d'informations concernant la durabilité sur le traitement perceptif/sensoriel : identifier comment le type de conceptualisation de la durabilité module la perception hédonique de différents types de produits.
- 3) Identification de leviers pour adopter une alimentation plus durable : à partir des résultats obtenus sur les deux premiers objectifs, le doctorant tentera d'identifier des leviers pour les comportements des consommateurs vers des choix alimentaires plus durables.

Enfin, comme on l'a vu, le profilage des drogues illicites a une longue histoire. Plusieurs pays ont déployé des efforts pour développer et mettre en œuvre l'utilisation systématique du profilage des drogues illicites dans le cadre des enquêtes et des poursuites relatives aux crimes liés aux drogues illicites. En France, les résultats du

profilage sont stockés, pour le moment, dans une base de données séparée de la **base de données nationale STUPS®** (Système de Traitement Uniformisé des Produits Stupéfiants) du ministère de l'Intérieur répertoriant les résultats d'analyses d'identification et de quantification (les 2 premières étapes du processus, la 3^e étant le profilage). Ces dernières contiennent certaines informations contextuelles. Plusieurs facteurs sont actuellement à l'origine de la perte de renseignements précieux et l'analyse et l'interprétation approfondies des données en font partie.

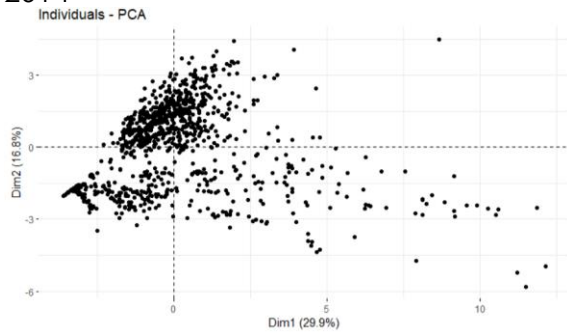
Ces bases contiennent aujourd'hui environ 10 millions d'entrées et présentent donc un vaste terrain de jeu pour une exploitation des données en vue d'extraire de la connaissance. Dans les domaines agronomiques et agroalimentaires, il n'existe pas d'équivalent. C'est pourquoi une collaboration s'est opérée entre le Service National de Police Scientifique (SNPS), le laboratoire d'informatique de Bourgogne (LIB) et la Cellule d'appui à la recherche en science des données de la Direction Scientifique de l'Institut Agro Dijon pour **déboucher sur le financement d'une thèse**.

Il s'agira de proposer un système intelligent pour répondre aux défis liés à l'interprétation de données hétérogènes et multivariées (modèles linéaires et non linéaires) contenues dans la base de données STUPS® afin d'en décrire, comprendre et expliquer les connaissances implicites. Les travaux de recherche visés dans cette thèse concernent le domaine de l'Intelligence Artificielle (IA), et s'orienteront sur deux aspects fondamentaux : IA symbolique (modèles de connaissances définissant les sémantiques et autres aspects symboliques permettant d'interpréter et de raisonner sur ces connaissances), d'une part, et, IA statistique (modèles d'apprentissage automatique de type réseaux artificiels de neurones permettant de construire des prédictions), d'autre part. Elle est programmée pour démarrer en septembre 2022.

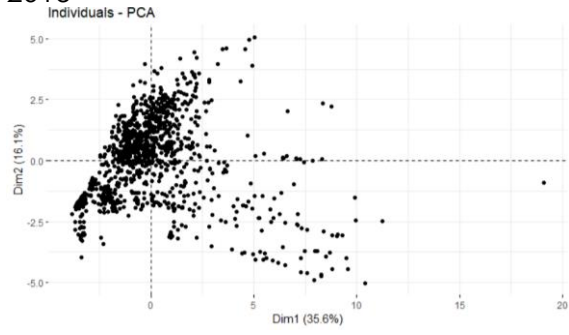
Une autre base (OTARIES pour Outil de Traitement Automatisé des Rapprochements Inter-Echantillons de Stupéfiants) regroupant les profils chimiques de la cocaïne et de la l'héroïne est alimentée par les laboratoires de Lyon et de Paris. Une harmonisation est en cours pour les 3 autres laboratoires (Lille, Toulouse et Marseille). Elle contient des milliers d'entrées sur plusieurs années. Les liens chimiques entre saisies sont établis à l'aide de méthodes de comparaison décrites dans le chapitre 2 (utilisation de la fonction cosinus), mais une étude plus globale pourrait être envisagée. Des méthodes multivariées sur des ensembles de profils regroupés par année ou sur une durée plus courte (à définir) peuvent être utilisées pour analyser les nuages de points et visualiser si d'une année sur l'autre le nuage est différent ou non. En effet, des clusters peuvent apparaître et d'autres disparaître. La figure 56 montre une analyse en composantes principales (ACP) des profils de solvants de la cocaïne par année, de 2014 à 2017.

[A]

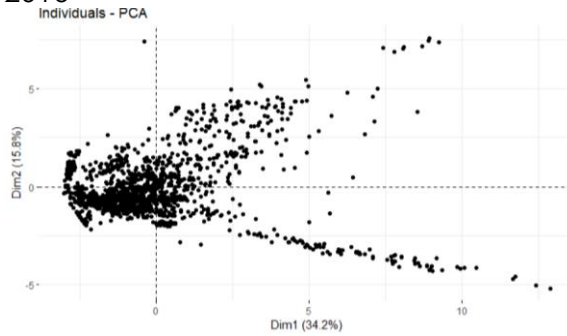
2014



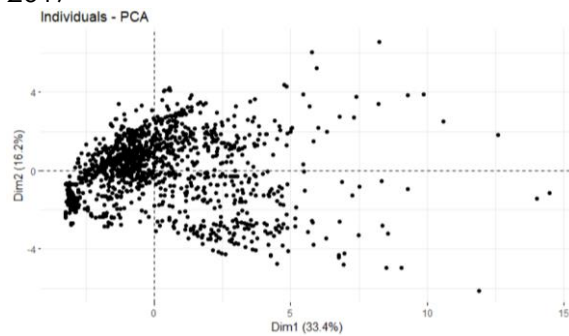
2015



2016

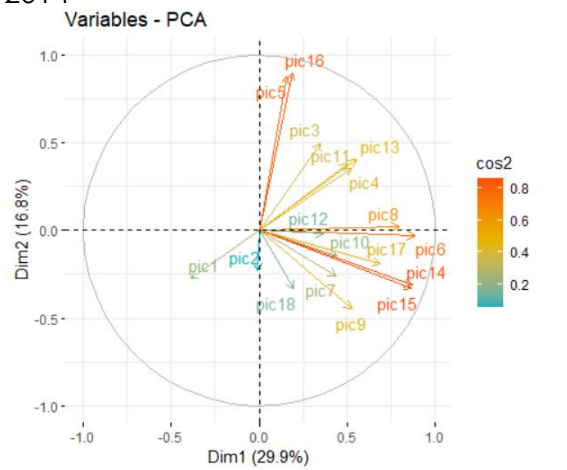


2017

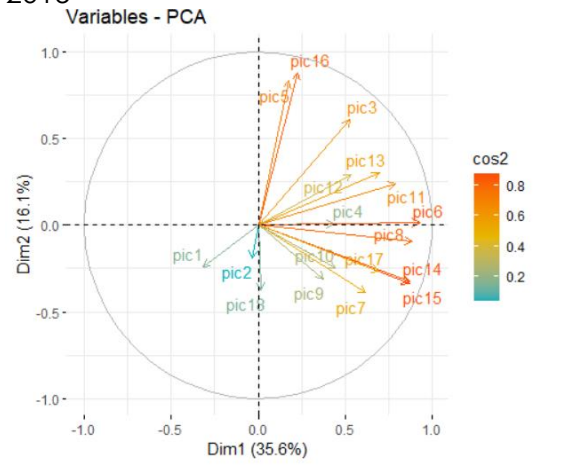


[B]

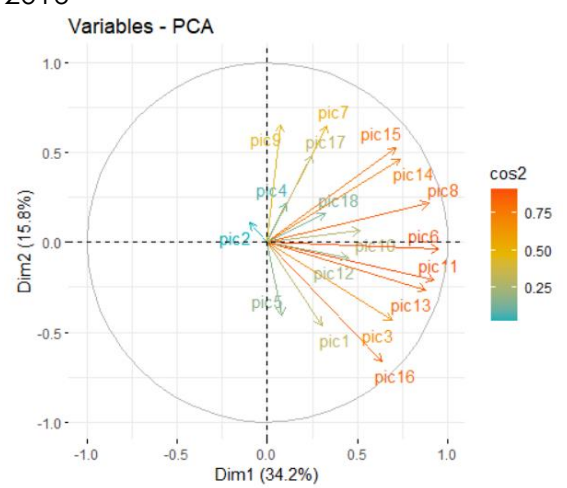
2014



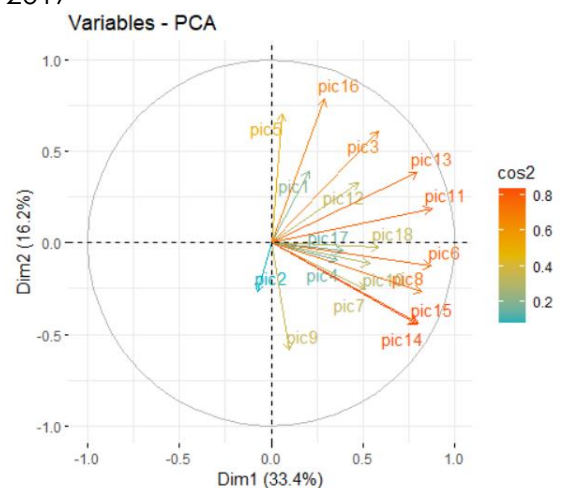
2015



2016



2017



[C]

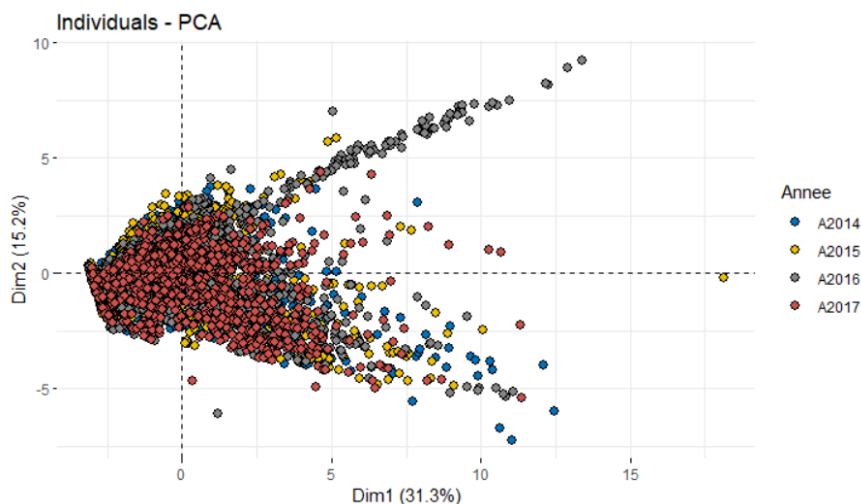


Figure 56 : Analyses en composantes principales des profils chimiques des solvants de la cocaïne. [A] graphique des individus, [B] graphiques des variables et [C] graphique des individus sur les 4 années.

L'ACP ne permet pas de distinguer clairement des groupes isolés. Sur les graphiques d'individus, on constate plutôt une sorte de continuum. La variance totale expliquée avec 2 composantes est de 50% environ.

Sur les graphiques des variables, on peut constater que les variables les mieux représentées et qui contribuent le plus sont les mêmes au cours de ces 4 années ; on a les pics 6, 8, 11, 13, 14 et 16. Il s'agit des solvants : méthyl-ter-butyl éther (MTBE), inconnu1, dichloroethane, cyclohexane, inconnu2 et N-propylacétate. À creuser ce que cela signifie au niveau de la fabrication de la cocaïne.

Avec une autre analyse de réduction de dimensions t-SNE (t-distributed Stochastic Neighbor Embedding), on observe une structure en « éponge » dont quelques groupes se distinguent notamment en 3 dimensions cf. figure 57 [B].

Cette analyse est basée sur un algorithme non linéaire d'extraction de caractéristiques qui construit une nouvelle représentation des données de telle sorte que les données proches dans l'espace original ont une probabilité élevée d'avoir des représentations proches dans le nouvel espace. À l'inverse, les données qui sont éloignées dans l'espace original ont une probabilité faible d'avoir des représentations proches dans le nouvel espace. En pratique la similarité entre chaque paire de données, dans les 2 espaces, est mesurée par le biais de calculs probabilistes basés sur des hypothèses de distribution. Et les nouvelles représentations se construisent de telle sorte à minimiser la différence entre les distributions de probabilités mesurées dans l'espace original et celles du nouvel espace¹¹⁵.

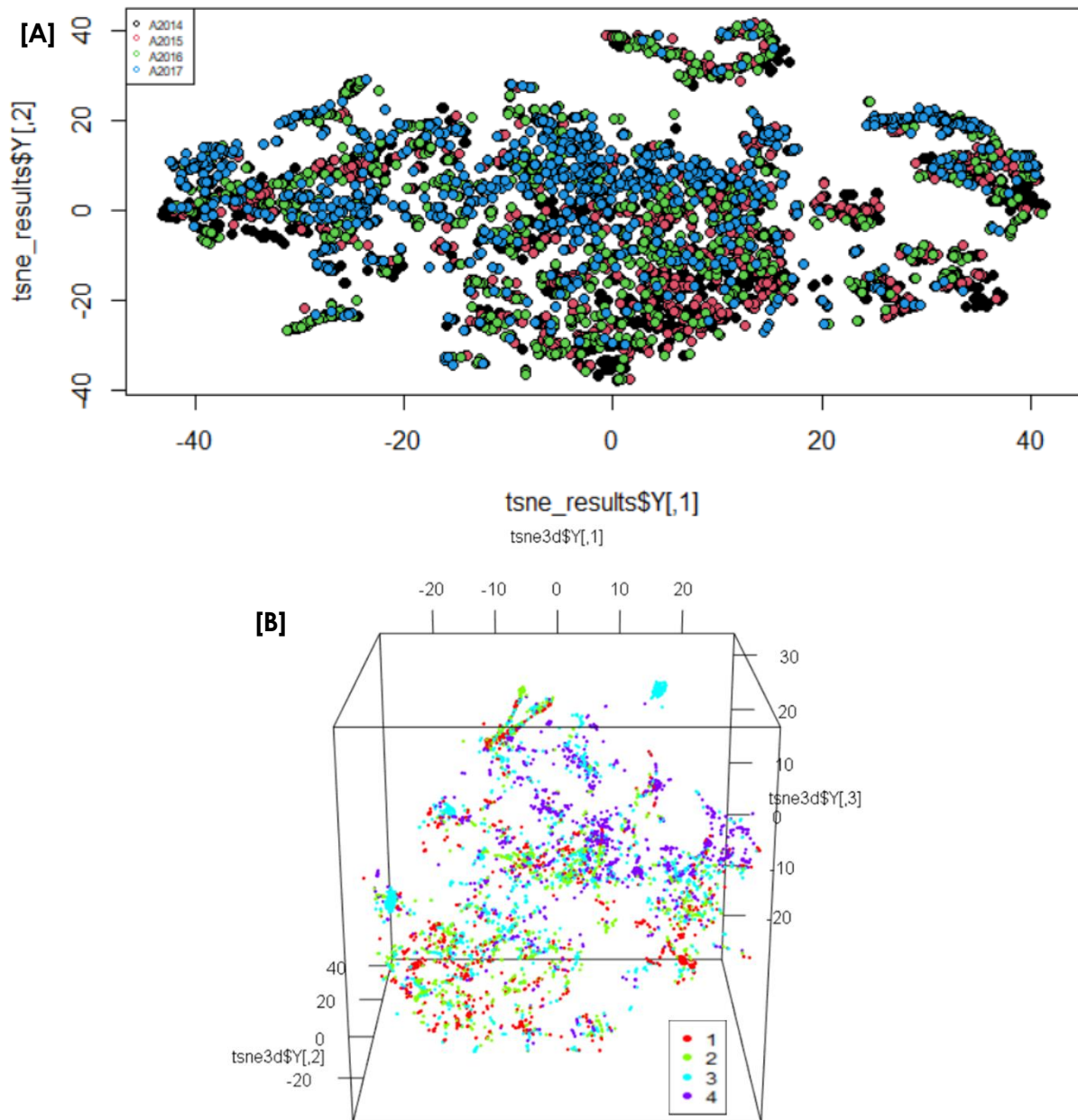


Figure 57 : Visualisation t-SNE. Graphique en 2 dimensions [A] et en 3D [B].

Suite à ces prémisses, j'aimerais explorer le potentiel de l'analyse de données topologiques (topological data analysis – TDA)¹¹⁶. Elle permet d'interagir avec des données structurées et non structurées et de les représenter par le biais d'un réseau topologique. En général, un réseau topologique fournit une carte de tous les points de l'ensemble de données, de sorte que les points proches sont plus similaires que les points éloignés. Le réseau est tout à fait analogue à une carte géographique, et joue le même rôle dans la compréhension du "paysage" des données.

C'est une méthode non supervisée (on ne fixe pas de classes ou groupes a priori).

La représentation compressée de la carte TDA présente plusieurs qualités :

- Les propriétés des données à petite et grande échelle sont préservées : ceci contraste avec d'autres méthodes de représentation qui déforment souvent les relations entre les points de données.

- Les données sont plus faciles à visualiser : les données réelles sont souvent de très haute dimension, et en tant que telles, elles sont difficiles à comprendre. Les cartes TDA, en revanche, sont faciles à comprendre.

Une grande partie de la TDA est basée sur la notion de proximité entre des points. La distance euclidienne peut être utilisée pour des données numériques, mais il existe, d'autres distances. Pour les données catégorielles, la notion de distance peut se faire soit en examinant les correspondances (c.-à-d. en définissant la distance entre les points par le nombre d'entrées qui sont identiques), soit en incluant d'autres mesures de similarité¹¹⁷.

C'est une méthode relativement récente qui est utilisée dans de nombreux domaines¹¹⁸ notamment les sciences de la vie^x (neuroscience, génomique, protéomique, biologie du cancer¹¹⁹, etc.).

L'étude de la structure des profils chimique a pour objectif de mettre en évidence des groupes particuliers qui apparaîtraient ou disparaîtraient au fil du temps. On obtiendrait ainsi une cartographie globale du réseau. Les modifications au cours du temps de la répartition des groupes dans la cartographie fonctionneraient comme des indicateurs des changements de procédés de fabrication de la drogue. La mise à jour dans un laps de temps court (3 à 6 mois, durée approximative de l'écoulement d'un lot) permettrait de caractériser rapidement ces changements et pourrait être utile d'un point stratégique aux services de police.

^x Special Issue Title: Applications of Topological Data Analysis in the Life Sciences - Website: <https://www.mdpi.com/si/73225> consulté le 24/09/2022

CONCLUSION

La rédaction de ce manuscrit m'a permis de faire le point sur mes travaux passés, en cours et sur les orientations vers lesquelles je souhaite me diriger. Il est toujours intéressant et satisfaisant intellectuellement de « faire un point » et de se projeter à court et à moyen terme. C'est en tout cas comme cela que j'aurai vécu la rédaction de ce travail d'HDR.

Mes travaux apportent une contribution dans le domaine de l'analyse de données (surtout d'origine physico-chimique). Ils m'ont surtout permis d'acquérir une vision globale sur la manière de résoudre un problème particulier de chimie analytique. Cette maîtrise de la chaîne complète me convient particulièrement, sachant que je suis consciente qu'il faut qu'il y ait aussi des chercheurs qui travaillent dans un domaine plus spécifique pour aider à lever différents verrous scientifiques présents dans certaines étapes de la résolution.

De par mon parcours, j'ai eu la chance de pouvoir appréhender les 2 facettes présentes dans la recherche que sont la recherche appliquée et la recherche académique. Il est intéressant de voir à quel moment et dans quel contexte, le passage d'une connaissance passe du registre académique à celui d'appliqué. Quel que soit le projet mené, il me paraît indispensable d'avoir une stratégie et une méthodologie adaptée en fonction de la finalité de celui-ci, de son contexte et des besoins exprimés. C'est pourquoi je souhaite continuer à pouvoir explorer les multiples facettes de l'analyse des données pour enrichir ces critères.

Références bibliographiques

- (1) Ribaux, O.; Margot, P. Science Forensique. *Dictionnaire de criminologie en ligne*.
- (2) Margot, P. Un Changement de Nom Dans La Continuité. *Revue internationale de criminologie et de police technique et scientifique* **1999**, 52 (1), 6–8.
- (3) Shah, S. R.; Patel, M. A.; Naik, M. V.; Pradhan, P. K.; Upadhyay, U. M. Recent Approaches of "impurity Profiling in Pharmaceutical Analysis: A Review. *International Journal of Pharmaceutical Sciences And Research* **2012**, 3 (10), 3603–3617.
- (4) Otto, M. *Chemometrics: Statistics and Computer Application in Analytical Chemistry*, 3rd ed.; Wiley VCH: Weinheim, 2017.
- (5) Esbensen, K.; Geladi, P. The Start and Early History of Chemometrics: Selected Interviews. Part 2. *Journal of Chemometrics* **1990**, 4 (6), 389–412. <https://doi.org/10.1002/cem.1180040604>.
- (6) Brereton, R. G. *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*; John Wiley & Sons, Ltd: Chichester, 2003.
- (7) Rivière, P. Qu'est-Ce Qu'une Donnée ? Impact Des Données Externes Sur La Statistique Publique. *Courier des statistiques* **2020**, No. 5, 114–131.
- (8) European Monitoring Centre for Drugs and Drug Addiction. *New Psychoactive Substances: 25 Years of Early Warning and Response in Europe. An Update from the EU Early Warning System (June 2022)*; EMCDDA; Publications Office of the European Union: Luxembourg, 2022; p 42.
- (9) Dujourdy, L.; Santin, B.; Soto, T. Chimie Des Couleurs et Investigations Préliminaires En Police Scientifique. *L'actualité chimique* **2013**, No. 378–379, 52–60.
- (10) Richard, D. *Dictionnaire des drogues, des toxicomanies et des dépendances*; Senon, J.-Louis., Kouchner, B., Series Eds.; Les référents; Larousse: Paris, 1999.
- (11) Dujourdy, L.; Besacier, F. L'héroïne Saisie En France. Données Statistiques Issues de La Base Nationale Des Laboratoires de Police Scientifique. *Annales Pharmaceutiques Françaises* **2010**, 68 (2), 127–132. <https://doi.org/10.1016/j.pharma.2010.02.002>.
- (12) Dujourdy, L.; Besacier, F. La Cocaïne Saisie En France. Exploitation Des Données Statistiques Nationales. *L'actualité chimique* **2010**, No. 342–343, 29–36.
- (13) Broséus, J.; Gentile, N.; Esseiva, P. The Cutting of Cocaine and Heroin: A Critical Review. *Forensic Science International* **2016**, 262, 73–83. <https://doi.org/10.1016/j.forsciint.2016.02.033>.
- (14) Fiorentin, T. R.; Krotulski, A. J.; Martin, D. M.; Browne, T.; Triplett, J.; Conti, T.; Logan, B. K. Detection of Cutting Agents in Drug-Positive Seized Exhibits within the United States. *Journal of Forensic Sciences* **2019**, 64 (3), 888–896. <https://doi.org/10.1111/1556-4029.13968>.
- (15) Dujourdy, L.; Besacier, F. A Study of Cannabis Potency in France over a 25 Years Period (1992–2016). *Forensic Science International* **2017**, 272, 72–80. <https://doi.org/10.1016/j.forsciint.2017.01.007>.
- (16) Esseiva, P.; Margot, P. Drug Profiling. *Wiley Encyclopedia of Forensic Science*; John Wiley & Sons, Ltd, 2009. <https://doi.org/10.1002/9780470061589.fsa406>.
- (17) United Nations Office on Drugs and Crime, U. *Guidelines on Representative Drug Sampling*; Manual; United Nations New York, 2009.
- (18) ENFSI Drugs Working Group DWG. *Guidelines on Sampling of Illicit Drugs for Qualitative Analysis*; European Network of Forensic Institutes DWG, 2012.

- (19) Smith, P. L. *A Primer for Sampling Solids, Liquids, and Gases: Based on the Seven Sampling Errors of Pierre Gy*; Series on Statistics and Applied Probability; ASA-SIAM: Philadelphia, 2006.
- (20) Gerville-Réache, L.; Couallier, V. Échantillon Représentatif (d'une Population Finie) : Définition Statistique et Propriétés, 2011.
- (21) Aitken, C. Sampling and Estimation of Quantities. *Wiley Encyclopaedia of Forensic Science*; John Wiley & Sons Inc.: United States, 2009; pp 2281–2291.
- (22) ENFSI Drugs Working Group DWG. *Guidelines on Sampling of Illicit Drugs for Quantitative Analysis*; European Network of Forensic Institutes DWG, 2014.
- (23) Gy, P. M. *Sampling for Analytical Purposes*, 1st ed.; John Wiley & Sons, Ltd: Chichester, 1998.
- (24) Pitard, F. F. *Pierre Gy's Sampling Theory and Sampling Practice*, Third.; CRC Press - Taylor & Francis Group: Boca Raton, 2019.
- (25) Dujourdy, L.; Csesztregi, T.; Bovens, M.; Franc, A.; Nagy, J. Sampling of Illicit Drugs for Quantitative Analysis. Part I: Heterogeneity Study of Illicit Drugs in Europe. *Forensic Science International* **2013**, 231 (1), 249–256. <https://doi.org/10.1016/j.forsciint.2013.05.022>.
- (26) Bovens, M.; Csesztregi, T.; Franc, A.; Nagy, J.; Dujourdy, L. Sampling of Illicit Drugs for Quantitative Analysis—Part II. Study of Particle Size and Its Influence on Mass Reduction. *Forensic Science International* **2014**, 234, 174–180. <https://doi.org/10.1016/j.forsciint.2013.09.001>.
- (27) Hubert, P.; Nguyen-Huu, J.-J.; Boulanger, B.; Chapuzet, E.; Chiap, P.; Cohen, N.; Compagnon, P. A.; Dewé, W.; Feinberg, M.; Lallier, M.; Laurentie, M.; Mercier, N.; Muzard, G.; Nivert, C.; Valat, L. Validation des procédures analytiques quantitatives. Harmonisation des démarches. *STP Pharma Pratiques* **2003**, 13 (3).
- (28) Hubert, P.; Nguyen-Huu, J.-J.; Boulanger, B.; Chapuzet, E.; Cohen, N.; Compagnon, P.-A.; Dewé, W.; Feinberg, M.; Laurentie, M.; Mercier, N.; Muzard, G. Validation des procédures analytiques quantitatives : Partie II - Statistiques. *STP Pharma Pratiques* **2006**, 16 (1).
- (29) Hubert, P.; Nguyen-Huu, J.-J.; Boulanger, B.; Chapuzet, E.; Cohen, N.; Compagnon, P.-A.; Dewé, W.; Feinberg, M.; Laurentie, M.; Mercier, N.; Muzard, G.; Valat, L. Validation des procédures analytiques quantitatives : harmonisation des démarches - Partie III. Exemples d'application. *STP Pharma Pratiques* **2006**, 16 (2).
- (30) Feinberg, M. Validation Interne Des Méthodes d'analyse. *Tech. de l'Ingénieur* **2001**, 224–228.
- (31) Dewé, W.; Govaerts, B.; Boulanger, B.; Rozet, E.; Chiap, P.; Hubert, P. Using Total Error as Decision Criterion in Analytical Method Transfer. *Chemometrics and Intelligent Laboratory Systems* **2007**, 85 (2), 262–268. <https://doi.org/10.1016/j.chemolab.2006.07.003>.
- (32) AFNOR. *NF ISO 5725-1 Application de La Statistique - Exactitude (Justesse et Fidélité) Des Résultats et Méthodes de Mesure - Partie 1 : Principes Généraux et Définitions*; 1994.
- (33) AFNOR, *5720-2. NF ISO 5720-2 Exactitude (Justesse et Fidélité) Des Résultats et Méthodes de Mesure - Partie 2 : Méthode de Base Pour La Détermination de La Répétabilité et de La Reproductibilité d'une Méthode de Mesure Normalisée*; 2020; Vol. NF ISO 5720-2.
- (34) AFNOR, *5725-3. NF ISO 5725-3 Application de La Statistique - Exactitude (Justesse et Fidélité) Des Résultats et Méthodes de Mesure - Partie 3 : Mesures*

- Intermédiaires de La Fidélité d'une Méthode de Mesure Normalisée*; 1994; Vol. NF ISO 5725-3.
- (35) AFNOR, 5725-4. *NF ISO 5725-4 Exactitude (Justesse et Fidélité) Des Résultats et Méthodes de Mesure - Partie 4 : Méthodes de Base Pour La Détermination de La Justesse d'une Méthode de Mesure Normalisée*; 2020; Vol. NF ISO 5725-4.
- (36) AFNOR, F. V.-116. *FD V03-116 Analyse Des Produits Agricoles et Alimentaires - Guide d'application Des Données Métrologiques*; 2001; Vol. FD V03-116.
- (37) R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021.
- (38) Rozet, E. Improvement of the Predictive Character of Test Results Issued from Analytical Methods Life Cycle. PhD Thesis, ULiège - Université de Liège, 2008.
- (39) Feinberg, M. *Labostat – Guide de Validation Des Méthodes d'analyse*; Lavoisier, 2009.
- (40) Dujourdy, L.; Charvoz, C.; Dalmaso, M.; Dufour, A.-B. À La Recherche Du Juste Dosage : Statistique Appliquée à l'analyse de La Cocaïne. *Annales Pharmaceutiques Françaises* **2013**, 71 (3), 193–200. <https://doi.org/10.1016/j.pharma.2013.01.005>.
- (41) Dujourdy, L.; Charvoz, C.; Dalmaso, M.; Dufour, A.-B. What a Validation Strategy Means for the Quantitation of Cocaine and Heroin? *Forensic Science International* **2015**, 251, 32–39. <https://doi.org/10.1016/j.forsciint.2015.03.009>.
- (42) Ladroue, V.; Dujourdy, L.; Besacier, F. Profilage Des Impuretés de l'amphétamine. **2010**, No. 342–343, 37–45.
- (43) Morelato, M.; Beavis, A.; Tahtouh, M.; Ribaux, O.; Kirkbride, P.; Roux, C. The Use of Forensic Case Data in Intelligence-Led Policing: The Example of Drug Profiling. *Forensic Science International* **2013**, 226 (1), 1–9. <https://doi.org/10.1016/j.forsciint.2013.01.003>.
- (44) Collins, M. Illicit Drug Profiling: A Historical Perspective. *Drug testing and analysis* **2021**.
- (45) United Nations Office on Drugs and Crime, U. *Methods for Impurity Profiling of Heroin and Cocaine*; Manual ST/NAR/35; UNODC: Vienna, Austria, 2005; p 81.
- (46) Sansfaçon, D.; Barchechat, O.; Lopez, D.; Valade, C. *Drogues et Dommages Sociaux - Revue de Littérature Internationale*; Rapport d'études; Observatoire Français des Drogues et Toxicomanies: Saint-Denis, 2005; p 456.
- (47) Guéniat, O. Le Profilage de l'héroïne et de La Cocaïne - Les Méthodes d'analyse, La Modélisation Du Concept Du Profilage, La Gestion et l'exploitation Des Liens, Lausanne, Institut de Police Scientifique et de Criminologie, 2001.
- (48) Besacier, F. Réalisation d'empreintes Isotopiques Pour La Recherche d'origine Des Drogues : Application à l'héroïne, Claude Bernard, Lyon I, 1999.
- (49) Bishop, C. M. *Pattern Recognition and Machine Learning*; New York : Springer, [2006] ©2006, 2006.
- (50) Hand, D. J.; Mannila, H.; Smyth, P. *Principles of Data Mining*, MIT Press.; Adaptive Computation and Machine Learning series; Cambridge, 2001.
- (51) Shirshorshidi, A. S.; Aghabozorgi, S.; Wah, T. Y. A Comparison Study on Similarity and Dissimilarity Measures in Clustering Continuous Data. *PLoS One* **2015**, 10 (12), e0144059–e0144059. <https://doi.org/10.1371/journal.pone.0144059>.
- (52) Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning – Data Mining, Inference and Prediction*; 2013.
- (53) Beebe, K. R.; Pell, R. J.; Seasholtz, M. B. *Chemometrics: A Practical Guide*; A Wiley-Interscience publication; Wiley, 1998.

- (54) Jonson, C. S. L.; Strömberg, L. Computer Aided Retrieval of Common-Batch Members in Leuckart Amphetamine Profiling. *Journal of forensic sciences* **1993**, *38* 6, 1472–1477.
- (55) Robertson, B.; Vignaux, G. A. *Interpreting Evidence: Evaluating Forensic Science in the Courtroom*; John Wiley & Sons, Ltd: Chichester, 1995.
- (56) Robertson, B.; Vignaux, G.; Berger, C. *Interpreting Evidence: Evaluating Forensic Science in the Courtroom*, 2nd ed.; John Wiley & Sons, Ltd: Chichester, 2016.
- (57) Zadora, G.; Martyna, A.; Ramos, D.; Aitken, C. G. G. *Statistical Analysis in Forensic Science: Evidential Value of Multivariate Physicochemical Data*, 1st ed.; John Wiley & Sons, Ltd: Chichester, 2014.
- (58) Curran, J. M.; Champod, T. N. H.; Buckleton, J. S. *Forensic Interpretation of Glass Evidence*, CRC Press LLC.; Taylor & Francis, 2000.
- (59) Ballany, J.; Caddy, B.; Cole, M.; Finnon, Y.; Aalberg, L.; Janhunen, K.; Sippola, E.; Andersson, K.; Bertler, C.; Dahlén, J.; Kopp, I.; Dujourdy, L.; Lock, E.; Margot, P.; Huizer, H.; Poortman, A.; Kaa, E.; Lopes, A. Development of a Harmonised Pan-European Method for the Profiling of Amphetamines. *Science & Justice* **2001**, *41* (3), 193–196. [https://doi.org/10.1016/S1355-0306\(01\)71889-8](https://doi.org/10.1016/S1355-0306(01)71889-8).
- (60) Aalberg, L.; Andersson, K.; Bertler, C.; Borén, H.; Cole, M. D.; Dahlén, J.; Finnon, Y.; Huizer, H.; Jalava, K.; Kaa, E.; Lock, E.; Lopes, A.; Meer, A. P. der; Sippola, E. Development of a Harmonised Method for the Profiling of Amphetamines: I. Synthesis of Standards and Compilation of Analytical Data. *Forensic Science International* **2005**, *149* (2), 219–229. <https://doi.org/10.1016/j.forsciint.2004.06.020>.
- (61) Aalberg, L.; Andersson, K.; Bertler, C.; Cole, M. D.; Finnon, Y.; Huizer, H.; Jalava, K.; Kaa, E.; Lock, E.; Lopes, A.; Meer, A. P. der; Sippola, E.; Dahlén, J. Development of a Harmonised Method for the Profiling of Amphetamines: II. Stability of Impurities in Organic Solvents. *Forensic Science International* **2005**, *149* (2), 231–241. <https://doi.org/10.1016/j.forsciint.2004.06.019>.
- (62) Andersson, K.; Jalava, K.; Lock, E.; Finnon, Y.; Huizer, H.; Kaa, E.; Lopes, A.; Meer, A. P. der; Cole, M. D.; Dahlén, J.; Sippola, E. Development of a Harmonised Method for the Profiling of Amphetamines: III. Development of the Gas Chromatographic Method. *Forensic Science International* **2007**, *169* (1), 50–63. <https://doi.org/10.1016/j.forsciint.2006.10.018>.
- (63) Andersson, K.; Jalava, K.; Lock, E.; Huizer, H.; Kaa, E.; Lopes, A.; Meer, A. P. der; Cole, M. D.; Dahlén, J.; Sippola, E. Development of a Harmonised Method for the Profiling of Amphetamines: IV. Optimisation of Sample Preparation. *Forensic Science International* **2007**, *169* (1), 64–76. <https://doi.org/10.1016/j.forsciint.2006.10.017>.
- (64) Lock, E.; Aalberg, L.; Andersson, K.; Dahlén, J.; Cole, M. D.; Finnon, Y.; Huizer, H.; Jalava, K.; Kaa, E.; Lopes, A.; Meer, A. P. der; Sippola, E. Development of a Harmonised Method for the Profiling of Amphetamines V: Determination of the Variability of the Optimised Method. *Forensic Science International* **2007**, *169* (1), 77–85. <https://doi.org/10.1016/j.forsciint.2006.10.019>.
- (65) Andersson, K.; Lock, E.; Jalava, K.; Huizer, H.; Jonson, S.; Kaa, E.; Lopes, A.; Meer, A. P. der; Sippola, E.; Dujourdy, L.; Dahlén, J. Development of a Harmonised Method for the Profiling of Amphetamines VI: Evaluation of Methods for Comparison of Amphetamine. *Forensic Science International* **2007**, *169* (1), 86–99. <https://doi.org/10.1016/j.forsciint.2006.10.020>.
- (66) Stojanovska, N.; Fu, S.; Tahtouh, M.; Kelly, T.; Beavis, A.; Kirkbride, K. P. A Review of Impurity Profiling and Synthetic Route of Manufacture of Methylamphetamine, 3,4-Methylenedioxymethylamphetamine,

- Amphetamine, Dimethylamphetamine and p-Methoxyamphetamine. *Forensic Science International* **2013**, 224 (1), 8–26. <https://doi.org/10.1016/j.forsciint.2012.10.040>.
- (67) United Nations Office on Drugs and Crime, U. *Recommended Methods for the Identification and Analysis of Amphetamine, Methamphetamine and Their Ring-Substituted Analogues in Seized Materials*; Manual ST/NAR/34; United Nations New York: Vienna, Austria, 2006; p 88.
- (68) Spadari, M.; Batisse, A.; Guerlais, M.; Boucher, A.; Daveluy, A.; Boisselier, R. L.; Gibaja, V.; Eiden, C.; Lepelley, M.; Roussin, A.; Deheul, S.; Frauger, E.; Debruyne, D. Les Aspects Actuels de l'usage de l'ecstasy/MDMA En France. *Thérapies* **2016**, 71 (6), 553–562. <https://doi.org/10.1016/j.therap.2016.05.005>.
- (69) Freudenmann, R. W.; Oxler, F.; Bernschneider-Reif, S. The Origin of MDMA (Ecstasy) Revisited: The True Story Reconstructed from the Original Documents. *Addiction* **2006**, 101 9, 1241–1245.
- (70) Heather, E.; Shimmon, R.; McDonagh, A. M. Organic Impurity Profiling of 3,4-Methylenedioxymethamphetamine (MDMA) Synthesised from Catechol. *Forensic Science International* **2015**, 248, 140–147. <https://doi.org/10.1016/j.forsciint.2014.12.021>.
- (71) Dujourdy, L.; Besacier, F.; Charvoz, C.; Ladroue, V. L'institut National de Police Scientifique, La Contribution d'un Service Répressif Sur l'axe Prévention. *La Lettre du Respadd (réseau de prévention des addictions)* **2015**, No. 22, 2–8.
- (72) Weyermann, C.; Marquis, R.; Delaporte, C.; Esseiva, P.; Lock, E.; Aalberg, L.; Bozenko, J. S.; Dieckmann, S.; Dujourdy, L.; Zrcek, F. Drug Intelligence Based on MDMA Tablets Data: I. Organic Impurities Profiling. *Forensic Science International* **2008**, 177 (1), 11–16. <https://doi.org/10.1016/j.forsciint.2007.10.001>.
- (73) Daré, B. L.; Morel, I.; Gicquel, T. Méthamphétamine : Où En Est-on 100 Ans Après Sa Découverte ? *La Presse Médicale Formation* **2020**, 1 (2), 222–228. <https://doi.org/10.1016/j.lpmfor.2020.04.014>.
- (74) Onoka, I.; Banyika, A. T.; Banerjee, P. N.; Makangara, J. J.; Dujourdy, L. A Review of the Newly Identified Impurity Profiles in Methamphetamine Seizures. *Forensic Science International: Synergy* **2020**, 2, 194–205. <https://doi.org/10.1016/j.fsisyn.2020.06.004>.
- (75) Dujourdy, L.; Dufey, V.; Besacier, F.; Miano, N.; Marquis, R.; Lock, E.; Aalberg, L.; Dieckmann, S.; Zrcek, F.; Bozenko, J. S. Drug Intelligence Based on Organic Impurities in Illicit MA Samples. *Forensic Science International* **2008**, 177 (2), 153–161. <https://doi.org/10.1016/j.forsciint.2007.11.013>.
- (76) Salouros, H. Illicit Drug Chemical Profiling: Current and Future State. *Australian Journal of Forensic Sciences* **2018**, 50, 689–696.
- (77) Onoka, I. Complementary Chemical Profiling of Methamphetamine for Strategic and Tactical Intelligence of Illicit Drugs, Dodoma, Tanzanie, 2021.
- (78) Tittarelli, R.; Luca, N. M. di; Pacifici, R.; Pichini, S.; Rio, A. D.; Busardò, F. P. Commentary - Heroin Purity and Adulteration: An Updated Snapshot from the Italian Early Warning System. *European review for medical and pharmacological sciences* **2020**, 24 8, 4461–4466.
- (79) United Nations Office on Drugs and Crime, U. *Méthodes Recommandées pour l'identification et l'analyse de la cocaïne contenue dans des substances Saisies (Texte Révisé et Mis à Jour) à l'usage des laboratoires nationaux d'analyse des drogues*; Manual ST/NAR/7/REV.1; UNODC: New York, 2012; p 48.
- (80) Dujourdy, L.; Taroni, F.; Margot, P. *Classification, interpretation and statistical evaluation of data in illicit drug profiling*; 1115-068017/2; Fonds National Suisse: Lausanne, 2003; p 74.

- (81) Esseiva, P.; Dujourdy, L.; Anglada, F.; Taroni, F.; Margot, P. A Methodology for Illicit Heroin Seizures Comparison in a Drug Intelligence Perspective Using Large Databases. *Forensic Science International* **2003**, *132* (2), 139–152. [https://doi.org/10.1016/S0379-0738\(03\)00010-0](https://doi.org/10.1016/S0379-0738(03)00010-0).
- (82) Esseiva, P.; Anglada, F.; Dujourdy, L.; Taroni, F.; Margot, P.; Pasquier, E. D.; Dawson, M.; Roux, C.; Doble, P. Chemical Profiling and Classification of Illicit Heroin by Principal Component Analysis, Calculation of Inter Sample Correlation and Artificial Neural Networks. *Talanta* **2005**, *67* (2), 360–367. <https://doi.org/10.1016/j.talanta.2005.03.041>.
- (83) Huttunen, J.; Austin, C.; Dawson, M.; Roux, C.; Robertson, J. Physical Evidence in Drug Intelligence, Part 1: Rationale Based on Hierarchic Distribution of Drugs Using Pyrolysis Gas Chromatography –Mass Spectrometry as an Example. *Australian Journal of Forensic Sciences* **2007**, *39* (2), 93–106. <https://doi.org/10.1080/00450610701650070>.
- (84) Locicero, S. Profilage de La Cocaine : Approche Inter-Laboratoires, Lausanne, Institut de Police Scientifique et de Criminologie, 2007.
- (85) Locicero, S.; Dujourdy, L.; Esseiva, P.; Besacier, F.; Hayoz, P.; Anglada, F.; Margot, P.; Guitton, J. Apport Scientifique à La Lutte Contre Le Phénomène Transfrontalier de Stupéfiants: Mise En Réseau de l'information. *Revue internationale de criminologie et de police technique et scientifique* **2005**, *LVIII* (1), 93–92.
- (86) Locicero, S.; Hayoz, P.; Esseiva, P.; Dujourdy, L.; Besacier, F.; Margot, P. Cocaine Profiling for Strategic Intelligence Purposes, a Cross-Border Project between France and Switzerland: Part I. Optimisation and Harmonisation of the Profiling Method. *Forensic Science International* **2007**, *167* (2), 220–228. <https://doi.org/10.1016/j.forsciint.2006.06.052>.
- (87) Locicero, S.; Esseiva, P.; Hayoz, P.; Dujourdy, L.; Besacier, F.; Margot, P. Cocaine Profiling for Strategic Intelligence, a Cross-Border Project between France and Switzerland: Part II. Validation of the Statistical Methodology for the Profiling of Cocaine. *Forensic Science International* **2008**, *177* (2), 199–206. <https://doi.org/10.1016/j.forsciint.2007.12.008>.
- (88) Dufey, V.; Dujourdy, L.; Besacier, F.; Chaudron, H. A Quick and Automated Method for Profiling Heroin Samples for Tactical Intelligence Purposes. *Forensic Science International* **2007**, *169* (2), 108–117. <https://doi.org/10.1016/j.forsciint.2006.08.003>.
- (89) Dujourdy, L.; Barbati, G.; Taroni, F.; Guéniat, O.; Esseiva, P.; Anglada, F.; Margot, P. Evaluation of Links in Heroin Seizures. *Forensic Science International* **2003**, *131* (2), 171–183. [https://doi.org/10.1016/S0379-0738\(02\)00432-2](https://doi.org/10.1016/S0379-0738(02)00432-2).
- (90) Bolck, A.; Weyermann, C.; Dujourdy, L.; Esseiva, P.; Berg, J. van den. Different Likelihood Ratio Approaches to Evaluate the Strength of Evidence of MDMA Tablet Comparisons. *Forensic Science International* **2009**, *191* (1), 42–51. <https://doi.org/10.1016/j.forsciint.2009.06.006>.
- (91) Brun, J.-D. La Lutte Contre Le Trafic de Cocaine Dans Les Caraïbes et En Afrique de l'Ouest. Un Enjeu Européen. Thèse de doctorat, Université des Antilles, Université des Antilles, 2019.
- (92) United Nations Office on Drugs and Crime, U. *Méthodes recommandées pour l'identification et l'analyse des cathinones de synthèse contenues dans des substances saisies Manuel à l'usage des laboratoires nationaux d'analyse des drogues*; Manual ST/NAR/49; New York, 2016; p 48.
- (93) European Monitoring Centre for Drugs and Drug Addiction, E. *EMCDDA INSIGHTS Cannabis Production and Markets in Europe*, Office for Official

- Publications of the European Communities.; EMCDDA INSIGHTS; Luxembourg, 2012.
- (94) Rossignol, J.; Dujourdy, L.; Stuerger, D.; Cayot, P.; Gougeon, R. D.; Bou-Maroun, E. A First Tentative for Simultaneous Detection of Fungicides in Model and Real Wines by Microwave Sensor Coupled to Molecularly Imprinted Sol-Gel Polymers. *Sensors (Basel)* **2020**, *20* (21), 6224. <https://doi.org/10.3390/s20216224>.
- (95) Daoud, S.; Bou-maroun, E.; Dujourdy, L.; Waschatko, G.; Billecke, N.; Cayot, P. Fast and Direct Analysis of Oxidation Levels of Oil-in-Water Emulsions Using ATR-FTIR. *Food Chemistry* **2019**, *293*, 307–314. <https://doi.org/10.1016/j.foodchem.2019.05.005>.
- (96) Lebbos, N.; Daou, C.; Ouaini, R.; Chebib, H.; Afram, M.; Curmi, P.; Dujourdy, L.; Bou-Maroun, E.; Chagnon, M.-C. Lebanese Population Exposure to Trace Elements via White Bread Consumption. *Foods* **2019**, *8* (11), 574. <https://doi.org/10.3390/foods8110574>.
- (97) Lebbos, N.; Keller, C.; Dujourdy, L.; Afram, M.; Curmi, P.; Darwish, T.; Daou, C.; Bou-Maroun, E. Validation of a New Method for Monitoring Trace Elements in Mediterranean Cereal Soils. *International Journal of Environmental Analytical Chemistry* **2021**, *0* (0), 1–17. <https://doi.org/10.1080/03067319.2021.1953000>.
- (98) Duley, G.; Dujourdy, L.; Klein, S.; Werwein, A.; Spartz, C.; Gougeon, R. D.; Taylor, D. K. Regionality in Australian Pinot Noir Wines: A Study on the Use of NMR and ICP-MS on Commercial Wines. *Food Chemistry* **2021**, *340*, 127906. <https://doi.org/10.1016/j.foodchem.2020.127906>.
- (99) Pagès-Hélary, S.; Dujourdy, L.; Cayot, N. Identification of Volatile Compounds in Blackcurrant Berries: Differences among Cultivars. *Molecules* **2021**, *26* (20), 6254. <https://doi.org/10.3390/molecules26206254>.
- (100) Pagès-Hélary, S.; Dujourdy, L.; Cayot, N. Flavor Compounds in Blackcurrant Berries: Multivariate Analysis of Datasets Obtained with Natural Variability and Various Experimental Parameters. *LWT* **2022**, *153*, 112425. <https://doi.org/10.1016/j.lwt.2021.112425>.
- (101) Daoud, S. Oxydation de Lipides En Présence de Fer : Aperçu Sur Le Rôle Du Tampon et Suivi Rapide Par Spectroscopie Infrarouge, Bourgogne, Dijon, 2020.
- (102) Lebbos, N. Tracking Trace Elements from Soil to Lebanese Bread and Populations Exposure. PhD Thesis, 2021.
- (103) Modugno, C.; Peltier, C.; Simonin, H.; Dujourdy, L.; Capitani, F.; Sandt, C.; Perrier-Cornet, J.-M. Understanding the Effects of High Pressure on Bacterial Spores Using Synchrotron Infrared Spectroscopy. *Frontiers in Microbiology* **2020**, *10*. <https://doi.org/10.3389/fmicb.2019.03122>.
- (104) Câmara, A. A.; Nguyen, T. D.; Saurel, R.; Sandt, C.; Peltier, C.; Dujourdy, L.; Husson, F. Biophysical Stress Responses of the Yeast *Lachancea Thermotolerans* During Dehydration Using Synchrotron-FTIR Microspectroscopy. *Frontiers in Microbiology* **2020**, *11*. <https://doi.org/10.3389/fmicb.2020.00899>.
- (105) Melendrez Ruiz, J. Understanding the Low Consumption of Pulses among French Non-Vegetarian Consumers: Combining Direct and Indirect Approaches to Identify Barriers and Opportunities. PhD Thesis, 2020.
- (106) Melendrez-Ruiz, J.; Goisbault, I.; Charrier, J.-C.; Pagnat, K.; Dujourdy, L.; Arvisenet, G.; Chambaron, S. An Exploratory Study Combining Eye-Tracking and Virtual Reality: Are Pulses Good “Eye-Catchers” in Virtual Supermarket Shelves? *Frontiers in Virtual Reality* **2021**, *2*. <https://doi.org/10.3389/frvir.2021.655273>.
- (107) Melendrez-Ruiz, J.; Dujourdy, L.; Goisbault, I.; Charrier, J.-C.; Pagnat, K.; Nicklaus, S.; Arvisenet, G.; Chambaron, S. “You Look at It, but Will You Choose It”: Is There

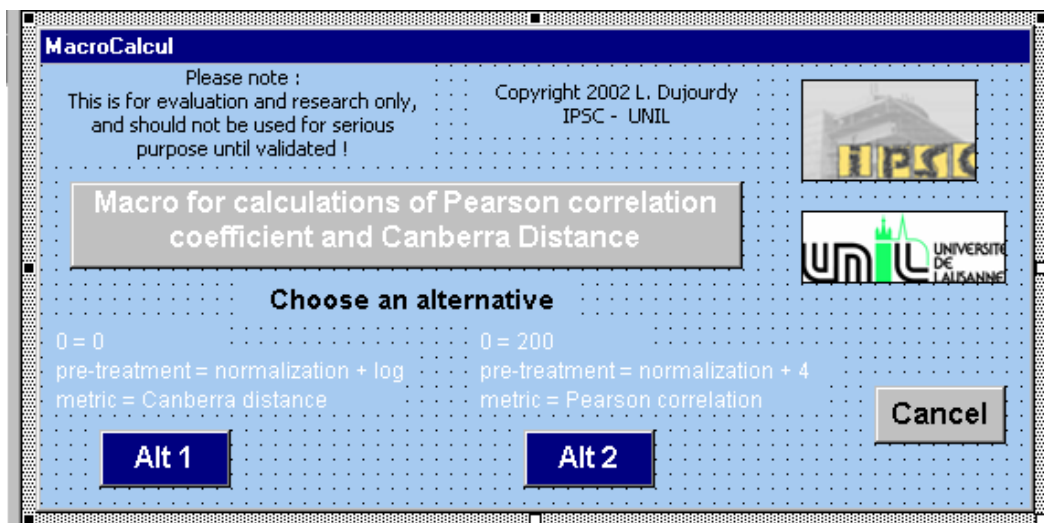
- a Link between the Foods Consumers Look at and What They Ultimately Choose in a Virtual Supermarket? *Food Quality and Preference* **2022**, 98, 104510. <https://doi.org/10.1016/j.foodqual.2021.104510>.
- (108) Husson, F.; Lê, S.; Pagès, J. *Analyse de Données Avec R*, 2e.; Pratique de la statistique; Presses Universitaires de Rennes, 2016.
- (109) Bouveyron, C.; Celeux, G.; Murphy, T. B.; Raftery, A. E. *Model-Based Clustering and Classification for Data Science: With Applications in R*; Cambridge Series in Statistical and Probabilistic Mathematics; Cambridge University Press: Cambridge, 2019. <https://doi.org/10.1017/9781108644181>.
- (110) Moscarola, J. *Faire parler les données. Méthodologies quantitatives et qualitatives*; Business Science Institute; EMS Editions: Caen, 2018.
- (111) Audigier, V.; Husson, F.; Josse, J. MIMCA: Multiple Imputation for Categorical Variables with Multiple Correspondence Analysis. *Statistics and Computing* **2017**, 27 (2), 501–518. <https://doi.org/10.1007/s11222-016-9635-4>.
- (112) Deneulin, P.; Bavaud, F. Analyses of Open-Ended Questions by Renormalized Associativities and Textual Networks: A Study of Perception of Minerality in Wine. *Food Quality and Preference* **2016**, 47, 34–44. <https://doi.org/10.1016/j.foodqual.2015.06.013>.
- (113) Deneulin, P.; Fur, Y. L.; Bavaud, F. Study of the Polysemic Term of Minerality in Wine: Segmentation of Consumers Based on Their Textual Responses to an Open-Ended Survey. *Food Research International* **2016**, 90, 288–297. <https://doi.org/10.1016/j.foodres.2016.11.004>.
- (114) Louis, P.-Y.; Robin, N. Une Mobilité d'une Extraordinaire Singularité : Les Mineur(e)s de l'Afrique Subsaharienne Aux Rives Sud de La Méditerranée. In *Etudier les Migr. Int. expériences Rech.*; 2018.
- (115) Maaten, L. van der; Hinton, G. E. Visualizing Data Using T-SNE. *Journal of Machine Learning Research* **2008**, 9, 2579–2605.
- (116) Carlsson, G. E. Topology and Data. *Bulletin of the American Mathematical Society* **2009**, 46, 255–308.
- (117) Boriah, S.; Chandola, V.; Kumar, V. Similarity Measures for Categorical Data: A Comparative Evaluation. In *Proceedings of the 2008 SIAM International Conference on Data Mining (SDM)*; 2008; pp 243–254. <https://doi.org/10.1137/1.9781611972788.22>.
- (118) Offroy, M.; Duponchel, L. Topological Data Analysis: A Promising Big Data Exploration Tool in Biology, Analytical Chemistry and Physical Chemistry. *Analytica Chimica Acta* **2016**, 910, 1–11. <https://doi.org/10.1016/j.aca.2015.12.037>.
- (119) Bukkuri, A.; Andor, N.; Darcy, I. K. Applications of Topological Data Analysis in Oncology. *Frontiers in Artificial Intelligence* **2021**, 4. <https://doi.org/10.3389/frai.2021.659037>.

ANNEXE 1

Explanations about construction of SMT macros in Excel® files:

SMTmacrocalculV1_021202.xls, SMTmacrocalculV2_021202.xls,
SMTmacrocalculV3_070103.xls and SMTmacrocalculV3b_070103.xls

- ◆ **First window** "Macrocalcul" which appears when you click on Execute macro button in CONDITIONS sheet.



- ◆ **Alt1 Button** corresponds to the alternative 1 macro :

Private Sub CommandButton3_Click() → Name of the macro corresponding to the clicking action on « button3 »

Dim nbsample As Integer → Variable definition

Dim sel As Variant

Dim selec As Variant

Dim plagesource As Variant

Dim plageÀremplir As Variant

Application.ScreenUpdating = False → Code for freezing screen during macro execution (time saving)

Sheets("Pretreat1").Select → Selection of « Pretreat1 » sheet

Range("a3:ah20000").Select → Selection of the range : « A3 » to « AH20000 »

Selection.ClearContents → Delete action

Sheets("board").Select → In « board » sheet

Range("ba1").FormulaR1C1 = "=COUNTA(C[-52])-2" → Calculation of number of samples and storage in « BA1 » cell.

nbsample = Range("ba1").Value → Allocation to « BA1 » value in nbsample integer variable

sel = "aa" & nbsample + 2 → sel variable corresponds to « AA » column with sample number +2 (e.g. if nbsample = 54 then sel is « AA56 »). It allows to know the last cell of data set

ANNEXE 1

selec = "aj" & nbsample + 2 → selec variable is « AJ » column with nbsample +2.
Corresponds to the last cell of Area sum column

Set plagesource = Worksheets("board").Range("aj2")
Set plageÀremplir = Worksheets("board").Range("aj2", selec)
plagesource.AutoFill Destination:=plageÀremplir → These commands are used to automatically get Area Sum of all target peaks from first to the last sample in « AJ » column

replacement → Call the « replacement » sub macro (see below)

Sheets("Pretreat1").Select → Go to « Pretreat1 » sheet
Set plagesource = Worksheets("Pretreat1").Range("b2:aa2")
Set plageÀremplir = Worksheets("Pretreat1").Range("b2", sel)
plagesource.AutoFill Destination:=plageÀremplir → Same command as above but for « Pretreat1 » sheet and range « B2 » to sel columns, sel is AA&nbsample+2 column. In this case, it automatically calculate for each cell the Log of normalised value (raw value/area sum of all target peaks)

Range("b3", sel).Select → Selection of line « B3 :sel »
Selection.Copy → Copy of the above selection
Sheets("board").Select → Selection of « board » sheet
Range("b3").Select → Selection of « B3 » cell
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False → Line « B3 :sel » in « Pretreat1 » sheet is copied in « B3 » cell in « board » sheet

Line2.Show → Show Line2 window (see below)
macrocalcul.Hide → Delete the first window : « macrocalcul »

End Sub → Code to stop the macro

- ◆ **Replacement macro** is used to replace cells with uncorrect values as "m", blank, space or 0 by Area Sum of all target peaks value for corresponding sample and allows to avoid 0 values in pretreatment with Log which give #NUMBER! error.

Sub replacement()

```
Dim nbsample As Integer
Dim selec As Variant
Dim selec2 As Variant
Dim i As Integer
Dim j As Integer
```

→ Recovery of number of samples (same as the first macro above)

```
Sheets("board").Select
Range("ba1").FormulaR1C1 = "=COUNTA(C[-52])-2"
nbsample = Range("ba1").Value
selec = "aj" & nbsample + 2
selec2 = "ak" & nbsample + 2
```

ANNEXE 1

Range("aj3", selec).Select → These 4 line program are used to preserve Sum of each variable values because when an erroneous cell will be replace by the sum of its variable. So column « AJ » is copied in column « AK »

```
Selection.Copy
Range("ak3").Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

For i = 3 To nbsample + 2 → First loop on overall samples

For j = 2 To 27 → Second loop on the 26 variables to research cells with incorrect values and correction by Area Sum of all target peaks for the sample in question

```
If Cells(i, j).Value = "" Then
Cells(i, j).Value = Cells(i, 37).Value
End If
If Cells(i, j).Value = " " Then
Cells(i, j).Value = Cells(i, 37).Value
End If
If Cells(i, j).Value = "m" Then
Cells(i, j).Value = Cells(i, 37).Value
End If
If IsEmpty(Cells(i, j)) Then
Cells(i, j).Value = Cells(i, 37).Value
End If
If Cells(i, j).Value = 0 Then
Cells(i, j).Value = Cells(i, 37).Value
End If
```

→ Replacement of erroneous value

Next j → Correction is done for the 1st sample (i=3 because the number of line is 3) and for all of its variables (j = 2 to 27 ; j= 2 because it corresponds to the 2nd column and the last column number corresponding to the last variable is 27) and then for the 2nd (i=4) and so on...

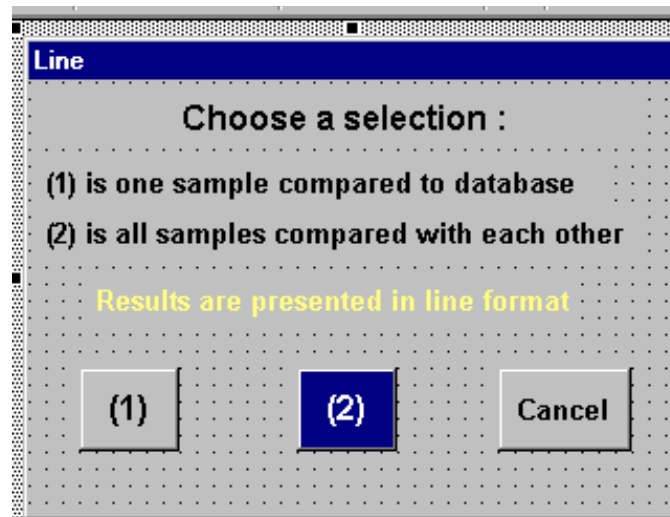
Next i

Range("ak3", selec2).Select → Reprocessing « AK » column values and copy in « AJ » column. Thus, Sum of each variable values correspond to the true values

```
Selection.Copy
Range("aj3").Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

End Sub

◆ **Second window** "Line2" which appears after the first macro.



- ◆ If you click on (1) button you access to the **“Onesample1 macro”** : is one sample compared to database

Sub Onesample1()

```
Dim nbsample As Integer
Dim nbsamp As Integer
Dim i As Integer
Dim j As Integer
Dim l1 As Integer
Dim l2 As Integer
Dim k As Double
Dim k2 As Double
Dim k3 As Double
Dim k4 As Double
Dim k5 As Double
Dim ll As Integer
Dim mm As Integer
Dim sel1 As Variant
Dim sel2 As Variant
Dim sel3 As Variant
Dim sel4 As Variant
Dim s1 As Variant
Dim s2 As Variant
Dim s3 As Variant
Dim s4 As Variant
Dim s5 As Variant
Dim selec As Variant
Dim plagesource As Variant
Dim plageÀremplir As Variant
```

Application.ScreenUpdating = False

Macro for the calculations of Canberra distances

```
Sheets("board").Select
Range("ba1").FormulaR1C1 = "=COUNTA(C[-52])-2"
nbsample = Range("ba1").Value
selec = "am" & nbsample + 2
```

```
Sheets("Calculine").Select
```

→ Data deleting in the previous sheets, to be sure that we work with new empty sheets

```
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine2").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine3").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine4").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine5").Select
Range("a2:g65536").Select
Selection.ClearContents
```

```
Sheets("transit").Select
Range("b2:ah3").Select
Selection.ClearContents
Range("b5:ah9").Select
Selection.ClearContents
```

```
k = 1
k2 = 1
k3 = 1
k4 = 1
k5 = 1
l1 = 3
nbsamp = nbsample
```

→ Allocation of macro variables

```
sel1 = "a" & 3
sel2 = "ah" & 3
```

```
For ll = 1 To nbsamp
```

→ First loop to calculate distances between sample of interest and the rest of the database. It first compares (i=1) sample 1 between itself and between sample 2... and so on, to nbsample

```
s1 = "a" & k + 1
s2 = "b" & k + 1
```

```
Sheets("transit").Select
```

→ Calculations are realised in « transit » sheet where the compararison is realised on each cell following Canberra formula. But before any calculation, it is necessary to delete all cells to avoid erroneous calculations

```
Range("b2:ah3").Select
Selection.ClearContents
Range("b5:ah9").Select
Selection.ClearContents
```

```
Sheets("Board").Select
Range(sel1, sel2).Select
Selection.Copy
Sheets("transit").Select
Range("a2").Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
Sheets("Board").Select
```

→ In « board », copy of line 2 and paste in « transit »

ANNEXE 1

```
Range(Cells(l1, 1), Cells(l1, 27)).Select      → In « board », copy of line l1 and paste in « transit »

Selection.Copy
Sheets("transit").Select
Range("a3").Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
l1 = l1 + 1                                     → l1 is incremented by 1 to compared the following
                                               database sample with sample of interest

For mm = 2 To 27                               → Second loop wich treats different cases

If IsEmpty(Cells(3, mm)) And Cells(2, mm).Value <> 0 Then
Cells(2, mm).Select
Selection.ClearContents
End If                                          → Numeric value compared with
                                               missing value is a non match and then
                                               numeric value is replaced by missing
                                               value

If IsEmpty(Cells(2, mm)) And Cells(3, mm).Value <> 0 Then
IsEmpty (Cells(9, mm))
Cells(3, mm).Select
Selection.ClearContents
End If                                          → Missing value compared with
                                               numeric value is a non match and
                                               then numeric value is replaced by
                                               missing value

If Cells(3, mm).Value <> 0 And Cells(2, mm).Value <> 0 Then
Cells(9, mm).Value = Abs(Cells(2, mm).Value - Cells(3, mm).Value) / Abs(Cells(2, mm).Value +
Cells(3, mm).Value)
End If                                          → Numeric value compared with
                                               numeric value is a match and then
                                               Canberra distance is calculated for
                                               the 2 compared cells. Value is
                                               assigned in cells(9,mm)
                                               corresponding to B9, C9, D9... etc
                                               when mm varies from 2 to 27

Next mm

Do While k < 65536                              → Third loop which copies value of total Canberra distance calculated
                                               in « AP5 » Column and paste it in « Calculine » sheet, in « B » column
                                               for each sample of the database. The total number of line in a Excel
                                               sheet is limited to 65536, so it is why there is this kind of loop

Sheets("transit").Select
Range("ap5").Select
Selection.Copy
Sheets("Calculine").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False

Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine").Select
```

ANNEXE 1

```
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k = k + 1
```

```
Exit Do
Loop
```

```
If k >= 65536 Then
```

→ Fourth loop, same as the third, to get the possibility to storage more than 65536 comparisons. Storage is done in « Calculine2 »

```
Do While k2 < 65536
```

```
s2 = "b" & k2 + 1
s1 = "a" & k2 + 1
Sheets("transit").Select
Range("ap5").Select
Selection.Copy
Sheets("Calculine2").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

```
Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine2").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k2 = k2 + 1
k = 0
```

```
Exit Do
Loop
End If
```

```
If k2 >= 65536 Then
```

→ Fifth loop, same as the third, to get the possibility to storage more than 131072 comparisons. Storage is done in « Calculine3 »

```
Do While k3 < 65536
```

```
s2 = "b" & k3 + 1
s1 = "a" & k3 + 1
Sheets("transit").Select
Range("ap5").Select
Selection.Copy
Sheets("Calculine3").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

```
Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine3").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k3 = k3 + 1
k = 0
```

ANNEXE 1

Exit Do
Loop
End If

If k3 >= 65536 Then

→ Sixth loop, same as the third, to get the possibility to storage more than 196608 comparisons. Storage is done in « Calculine4 »

Do While k4 < 65536

```
s2 = "b" & k4 + 1
s1 = "a" & k4 + 1
Sheets("transit").Select
Range("ap5").Select
Selection.Copy
Sheets("Calculine4").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

```
Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine4").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k4 = k4 + 1
k = 0
```

Exit Do
Loop
End If

If k4 >= 65536 Then

→ Seventh loop, same as the third, to get the possibility to storage more than 1048576 comparisons. Storage is done in « Calculine5 »

Do While k5 < 65536

```
s2 = "b" & k5 + 1
s1 = "a" & k5 + 1
Sheets("transit").Select
Range("ap5").Select
Selection.Copy
Sheets("Calculine5").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

```
Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine5").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k5 = k5 + 1
k = 0
```

Exit Do

ANNEXE 1

```
Loop  
End If
```

```
Next II
```

```
nbsamp = nbsamp - 1  
I1 = i + 3
```

```
End Sub
```

- ◆ If you click on (2) button you access to the **“Allsample1” macro** : is all samples compared to the other.

These Macro is based on the previous macro, the only change is the loop on the sample number.

```
Sub Allsample1()
```

```
Dim nbsample As Integer  
Dim nbsamp As Integer  
Dim i As Integer  
Dim j As Integer  
Dim I1 As Integer  
Dim I2 As Integer  
Dim k As Double  
Dim k2 As Double  
Dim k3 As Double  
Dim k4 As Double  
Dim k5 As Double  
Dim II As Integer  
Dim mm As Integer  
Dim sel1 As Variant  
Dim sel2 As Variant  
Dim sel3 As Variant  
Dim sel4 As Variant  
Dim s1 As Variant  
Dim s2 As Variant  
Dim s3 As Variant  
Dim s4 As Variant  
Dim s5 As Variant  
Dim selec As Variant  
Dim plagesource As Variant  
Dim plageÀremplir As Variant
```

```
Application.ScreenUpdating = False
```

```
'macro for calculations of Distance and Pearson correlation coefficient
```

```
Sheets("board").Select  
Range("ba1").FormulaR1C1 = "=COUNTA(C[-52])-2"  
nbsample = Range("ba1").Value  
selec = "am" & nbsample + 2
```

Data deleting in the previous sheets

```
Sheets("Calculine").Select  
Range("a2:g65536").Select  
Selection.ClearContents
```


ANNEXE 1

```
Sheets("Calculine2").Select  
Range("a2:g65536").Select  
Selection.ClearContents  
Sheets("Calculine3").Select  
Range("a2:g65536").Select  
Selection.ClearContents  
Sheets("Calculine4").Select  
Range("a2:g65536").Select  
Selection.ClearContents  
Sheets("Calculine5").Select  
Range("a2:g65536").Select  
Selection.ClearContents
```

```
Sheets("transit").Select  
Range("b2:ah3").Select  
Selection.ClearContents  
Range("b5:ah9").Select  
Selection.ClearContents
```

```
k = 1  
k2 = 1  
k3 = 1  
k4 = 1  
k5 = 1  
l1 = 3  
nbsamp = nbsample
```

```
For i = 1 To nbsample
```

! → First loop taking into account all of database samples for the calculations

```
sel1 = "a" & i + 2  
sel2 = "ah" & i + 2
```

```
Sheets("Board").Select  
Range(sel1, sel2).Select  
Selection.Copy  
Range("a2").Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False
```

```
For ll = 1 To nbsamp
```

```
s1 = "a" & k + 1  
s2 = "b" & k + 1
```

```
Sheets("transit").Select  
Range("b2:ah3").Select  
Selection.ClearContents  
Range("b5:ah9").Select  
Selection.ClearContents
```

```
Sheets("Board").Select  
Range(sel1, sel2).Select  
Selection.Copy  
Sheets("transit").Select  
Range("a2").Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
```

ANNEXE 1

```
False, Transpose:=False
Sheets("Board").Select
Range(Cells(11, 1), Cells(11, 27)).Select
Selection.Copy
Sheets("transit").Select
Range("a3").Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
l1 = l1 + 1

For mm = 2 To 27

If IsEmpty(Cells(3, mm)) And Cells(2, mm).Value <> 0 Then
IsEmpty (Cells(9, mm))
Cells(2, mm).Select
Selection.ClearContents
End If
If IsEmpty(Cells(2, mm)) And Cells(3, mm).Value <> 0 Then

IsEmpty (Cells(9, mm))
Cells(3, mm).Select
Selection.ClearContents
End If
If Cells(3, mm).Value <> 0 And Cells(2, mm).Value <> 0 Then
Cells(9, mm).Value = Abs(Cells(2, mm).Value - Cells(3, mm).Value) / Abs(Cells(2, mm).Value +
Cells(3, mm).Value)
End If
Next mm

Do While k < 65536

Sheets("transit").Select
Range("ap5").Select
Selection.Copy
Sheets("Calculine").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False

Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k = k + 1

Exit Do
Loop

If k >= 65536 Then
Do While k2 < 65536

s2 = "b" & k2 + 1
s1 = "a" & k2 + 1
Sheets("transit").Select
Range("ap5").Select
Selection.Copy
Sheets("Calculine2").Select
```

ANNEXE 1

```
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

```
Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine2").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k2 = k2 + 1
k = 0
```

```
Exit Do
Loop
End If
```

```
If k2 >= 65536 Then
Do While k3 < 65536
```

```
s2 = "b" & k3 + 1
s1 = "a" & k3 + 1
Sheets("transit").Select
Range("ap5").Select
Selection.Copy
Sheets("Calculine3").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

```
Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine3").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k3 = k3 + 1
k = 0
```

```
Exit Do
Loop
End If
```

```
If k3 >= 65536 Then
Do While k4 < 65536
```

```
s2 = "b" & k4 + 1
s1 = "a" & k4 + 1
Sheets("transit").Select
Range("ap5").Select
Selection.Copy
Sheets("Calculine4").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

```
Sheets("transit").Select
Range("a3").Select
```

ANNEXE 1

```
Selection.Copy
Sheets("calculine4").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k4 = k4 + 1
k = 0

Exit Do
Loop
End If

If k4 >= 65536 Then
Do While k5 < 65536

s2 = "b" & k5 + 1
s1 = "a" & k5 + 1
Sheets("transit").Select
Range("ap5").Select
Selection.Copy
Sheets("Calculine5").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False

Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine5").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k5 = k5 + 1
k = 0

Exit Do
Loop
End If

Next ll

nbsamp = nbsamp - 1
l1 = i + 3

Next i

End Sub
```

◆ **A1f2 Button** corresponds to the alternative 2 macro :

```
Private Sub CommandButton1_Click()           → Name of the macro corresponding to the clicking
                                              action on « button1 »
```

In this macro, missing or 0 values are replaced by 200

```
Dim nbsample As Integer
Dim Nbp As Integer
Dim cp As Integer
Dim i As Integer
```

ANNEXE 1

Dim j As Integer

Application.ScreenUpdating = False

Sheets("board").Select
nbsample = Range("ba1").Value
Nbp = Range("bb1").Value

→ Variable which recovers number of peaks with a numeric value

→ counter

cp = 26
Range("aj3,am10000").Select
Selection.ClearContents

For i = 3 To nbsample + 2

→ First loop on the sample overall

For j = 2 To 27

→ Second loop on target peaks. Replacement by 200 in cells containing missing, 0, or uncorrect values

If Cells(i, j).Value = "" Then
Cells(i, j).Value = 200
cp = cp - 1
End If
If Cells(i, j).Value = " " Then
Cells(i, j).Value = 200
cp = cp - 1
End If
If Cells(i, j).Value = "m" Then
Cells(i, j).Value = 200
cp = cp - 1
End If
If Cells(i, j).Value = 0 Then
Cells(i, j).Value = 200
cp = cp - 1
End If
Next j
Cells(i, 39).Value = cp

→ Recovering of real number of target peaks in « AM » column : « Nb peaks without 0 or missing peaks »

→ Counter is put on 26

cp = 26
Next i

pretreatment2
Line.Show
macrocalcul.Hide

→ Call sub macro « pretreatment2 » (see below)

→ Show Line window (see below)

→ Delete the first window : « macrocalcul »

End Sub

- ◆ **Pretreatment2 macro** is used to calculate pretreatment « normalization + 4th root », it is the same principle as Button3 macro for Alt1.

Sub pretreatment2()

Dim nbsample As Integer
Dim sel As Variant
Dim plagesource As Variant
Dim plageÀremplir As Variant

Application.ScreenUpdating = False

```
Sheets("Pretreat2").Select  
Range("a3:ah20000").Select  
Selection.ClearContents
```

```
Sheets("board").Select  
nbsample = Range("ba1").Value  
sel = "aa" & nbsample + 2  
selec = "aj" & nbsample + 2
```

```
Set plagesource = Worksheets("board").Range("aj2")  
Set plageÀremplir = Worksheets("board").Range("aj2", selec)  
plagesource.AutoFill Destination:=plageÀremplir
```

→ These commands are used to automatically get Area Sum of all target peaks from first to the last sample in « AJ » column

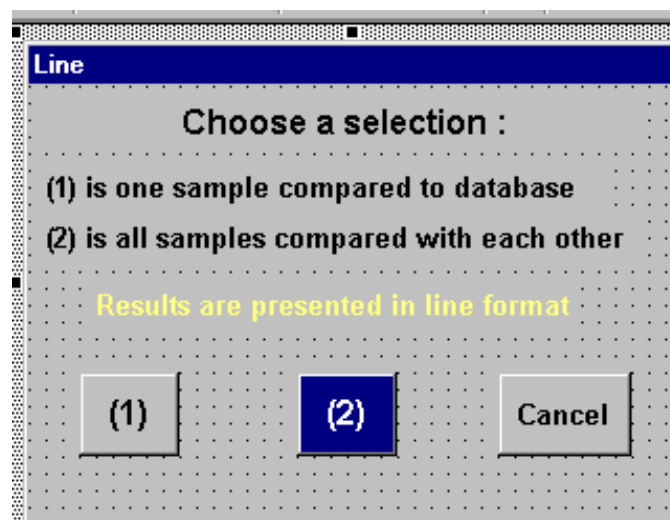
```
Sheets("Pretreat2").Select  
Set plagesource = Worksheets("Pretreat2").Range("b2:aa2")  
Set plageÀremplir = Worksheets("Pretreat2").Range("b2", sel)  
plagesource.AutoFill Destination:=plageÀremplir
```

→ Same command as above but for « Pretreat2 » sheet and range « B2 » to sel columns, sel is AA&nbsample+2 column. In this case, it automatically calculate for each cell the 4th root of normalised value (raw value/area sum of all target peaks)

```
Range("b3", sel).Select  
Selection.Copy  
Sheets("board").Select  
Range("b3").Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False
```

End Sub

- ◆ **Second window** "Line" which appears after this macro.



ANNEXE 1

- ◆ If you click on (1) button you access to the **"Onesample2 macro"** : is one sample compared to database

Sub Onesample2()

```
Dim nbsample As Integer
Dim nbsamp As Integer
Dim i As Integer
Dim j As Integer
Dim l1 As Integer
Dim l2 As Integer
Dim k As Double
Dim k2 As Double
Dim k3 As Double
Dim k4 As Double
Dim k5 As Double
Dim ll As Integer
Dim mm As Integer
Dim sel1 As Variant
Dim sel2 As Variant
Dim sel3 As Variant
Dim sel4 As Variant
Dim s1 As Variant
Dim s2 As Variant
Dim s3 As Variant
Dim s4 As Variant
Dim s5 As Variant
Dim selec As Variant
Dim plagesource As Variant
Dim plageÀremplir As Variant
```

```
Application.ScreenUpdating = False
```

```
Sheets("board").Select
Range("ba1").FormulaR1C1 = "=COUNTA(C[-52])-2"
nbsample = Range("ba1").Value
selec = "am" & nbsample + 2
```

Data deleting in the previous sheets

```
Sheets("Calculine").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine2").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine3").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine4").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine5").Select
Range("a2:g65536").Select
Selection.ClearContents
```

```
Sheets("transit").Select
Range("b2:ah3").Select
Selection.ClearContents
Range("b5:ah9").Select
```

ANNEXE 1

Selection.ClearContents

```
k = 1
k2 = 1
k3 = 1
k4 = 1
k5 = 1
l1 = 3
nbsamp = nbsample
```

```
sel1 = "a" & 3
sel2 = "ah" & 3
```

For ll = 1 To nbsamp

```
s1 = "a" & k + 1
s2 = "c" & k + 1
```

```
Sheets("transit").Select
Range("b2:ah3").Select
Selection.ClearContents
Range("b5:ah9").Select
Selection.ClearContents
```

```
Sheets("Board").Select
Range(sel1, sel2).Select
Selection.Copy
Sheets("transit").Select
Range("a2").Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
Sheets("Board").Select
Range(Cells(l1, 1), Cells(l1, 27)).Select
Selection.Copy
Sheets("transit").Select
Range("a3").Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
l1 = l1 + 1
```

For mm = 2 To 27

```
If IsEmpty(Cells(3, mm)) And Cells(2, mm).Value <> 0 Then
Cells(2, mm).Select
Selection.ClearContents
End If
If IsEmpty(Cells(2, mm)) And Cells(3, mm).Value <> 0 Then
Cells(3, mm).Select
Selection.ClearContents
End If
```

Next mm

Do While k < 65536

```
Sheets("transit").Select
Range("aq5").Select
Selection.Copy
Sheets("Calculine").Select
Range(s2).Select
```


ANNEXE 1

```
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False
```

```
Sheets("transit").Select  
Range("a3").Select  
Selection.Copy  
Sheets("calculine").Select  
Range(s1).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False  
k = k + 1
```

```
Exit Do  
Loop
```

```
If k >= 65536 Then  
Do While k2 < 65536
```

```
s2 = "c" & k2 + 1  
s1 = "a" & k2 + 1  
Sheets("transit").Select  
Range("aq5").Select  
Selection.Copy  
Sheets("Calculine2").Select  
Range(s2).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False
```

```
Sheets("transit").Select  
Range("a3").Select  
Selection.Copy  
Sheets("calculine2").Select  
Range(s1).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False  
k2 = k2 + 1  
k = 0
```

```
Exit Do  
Loop  
End If
```

```
If k2 >= 65536 Then  
Do While k3 < 65536
```

```
s2 = "c" & k3 + 1  
s1 = "a" & k3 + 1  
Sheets("transit").Select  
Range("aq5").Select  
Selection.Copy  
Sheets("Calculine3").Select  
Range(s2).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False
```

```
Sheets("transit").Select  
Range("a3").Select  
Selection.Copy  
Sheets("calculine3").Select  
Range(s1).Select
```

ANNEXE 1

```
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False  
k3 = k3 + 1  
k = 0
```

```
Exit Do  
Loop  
End If
```

```
If k3 >= 65536 Then  
Do While k4 < 65536
```

```
s2 = "c" & k4 + 1  
s1 = "a" & k4 + 1  
Sheets("transit").Select  
Range("aq5").Select  
Selection.Copy  
Sheets("Calculine4").Select  
Range(s2).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False
```

```
Sheets("transit").Select  
Range("a3").Select  
Selection.Copy  
Sheets("calculine4").Select  
Range(s1).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False  
k4 = k4 + 1  
k = 0
```

```
Exit Do  
Loop  
End If
```

```
If k4 >= 65536 Then  
Do While k5 < 65536
```

```
s2 = "c" & k5 + 1  
s1 = "a" & k5 + 1  
Sheets("transit").Select  
Range("aq5").Select  
Selection.Copy  
Sheets("Calculine5").Select  
Range(s2).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False
```

```
Sheets("transit").Select  
Range("a3").Select  
Selection.Copy  
Sheets("calculine5").Select  
Range(s1).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False  
k5 = k5 + 1  
k = 0
```

```
Exit Do
```

ANNEXE 1

```
Loop  
End If
```

```
Next II
```

```
nbsamp = nbsamp - 1  
I1 = i + 3
```

```
Sheets("board").Select
```

→ These lines are used to copy real number of peaks take into account for Pearson correlations between sample of interest and the the rest of database from « board » sheet to « Calculine » sheet

```
Range("am3", selec).Select
```

```
Selection.Copy
```

```
Sheets("calculine").Select
```

```
Range("e2").Select
```

```
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
```

```
False, Transpose:=False
```

```
End Sub
```

- ◆ If you click on (2) button you access to the **"Allsample2" macro** : is all samples compared to the other.

These Macro is also based on this previous macro, the only change is the loop on the sample number and recovery of number of compared peaks.

```
Sub Allsample2()
```

```
Dim nbsample As Integer
```

```
Dim nbsamp As Integer
```

```
Dim i As Integer
```

```
Dim j As Integer
```

```
Dim jj As Integer
```

```
Dim I1 As Integer
```

```
Dim I2 As Integer
```

```
Dim ii As Integer
```

```
Dim k As Double
```

```
Dim k2 As Double
```

```
Dim k3 As Double
```

```
Dim k4 As Double
```

```
Dim k5 As Double
```

```
Dim II As Integer
```

```
Dim mm As Integer
```

```
Dim sel1 As Variant
```

```
Dim sel2 As Variant
```

```
Dim sel3 As Variant
```

```
Dim sel4 As Variant
```

```
Dim s1 As Variant
```

```
Dim s2 As Variant
```

```
Dim s3 As Variant
```

```
Dim s4 As Variant
```

```
Dim s5 As Variant
```

```
Dim plagesource As Variant
```

```
Dim plageÀremplir As Variant
```

```
Application.ScreenUpdating = False
```

Macro for calculations of Pearson correlation coefficients

ANNEXE 1

```
Sheets("board").Select
Range("ba1").FormulaR1C1 = "=COUNTA(C[-52])-2"
nbsample = Range("ba1").Value
```

Data deleting in the previous sheets

```
Sheets("Calculine").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine2").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine3").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine4").Select
Range("a2:g65536").Select
Selection.ClearContents
Sheets("Calculine5").Select
Range("a2:g65536").Select
Selection.ClearContents
```

```
Sheets("transit").Select
Range("b2:ah3").Select
Selection.ClearContents
Range("b5:ah9").Select
Selection.ClearContents
```

```
k = 1
k2 = 1
k3 = 1
k4 = 1
k5 = 1
l1 = 3
jj = 2
ii = 3
nbsamp = nbsample
```

```
For i = 1 To nbsample
```

! → First loop taking into account all of database samples for the calculations

```
sel1 = "a" & i + 2
sel2 = "ah" & i + 2
```

```
Sheets("Board").Select
Range(sel1, sel2).Select
Selection.Copy
Range("a2").Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

Calculations of correlation coefficients

```
For ll = 1 To nbsamp
```

```
s1 = "a" & k + 1
s2 = "c" & k + 1
```

```
Sheets("transit").Select
```

ANNEXE 1

```
Range("b2:ah3").Select
Selection.ClearContents
Range("b5:ah9").Select
Selection.ClearContents

Sheets("Board").Select
Range(sel1, sel2).Select
Selection.Copy
Sheets("transit").Select
Range("a2").Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
Sheets("Board").Select
Range(Cells(l1, 1), Cells(l1, 27)).Select
Selection.Copy
Sheets("transit").Select
Range("a3").Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
l1 = l1 + 1

For mm = 2 To 27

If IsEmpty(Cells(3, mm)) And Cells(2, mm).Value <> 0 Then
Cells(2, mm).Select
Selection.ClearContents
End If
If IsEmpty(Cells(2, mm)) And Cells(3, mm).Value <> 0 Then
Cells(3, mm).Select
Selection.ClearContents
End If

Next mm

Do While k < 65536

Sheets("transit").Select
Range("aq5").Select
Selection.Copy
Sheets("Calculine").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False

Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k = k + 1

Exit Do
Loop

If k >= 65536 Then
Do While k2 < 65536

s2 = "c" & k2 + 1
```

ANNEXE 1

```
s1 = "a" & k2 + 1
Sheets("transit").Select
Range("aq5").Select
Selection.Copy
Sheets("Calculine2").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

```
Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine2").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k2 = k2 + 1
k = 0
```

```
Exit Do
Loop
End If
```

```
If k2 >= 65536 Then
Do While k3 < 65536
```

```
s2 = "c" & k3 + 1
s1 = "a" & k3 + 1
Sheets("transit").Select
Range("aq5").Select
Selection.Copy
Sheets("Calculine3").Select
Range(s2).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
```

```
Sheets("transit").Select
Range("a3").Select
Selection.Copy
Sheets("calculine3").Select
Range(s1).Select
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _
False, Transpose:=False
k3 = k3 + 1
k = 0
```

```
Exit Do
Loop
End If
```

```
If k3 >= 65536 Then
Do While k4 < 65536
```

```
s2 = "c" & k4 + 1
s1 = "a" & k4 + 1
Sheets("transit").Select
Range("aq5").Select
Selection.Copy
Sheets("Calculine4").Select
Range(s2).Select
```

ANNEXE 1

```
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False
```

```
Sheets("transit").Select  
Range("a3").Select  
Selection.Copy  
Sheets("calculine4").Select  
Range(s1).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False  
k4 = k4 + 1  
k = 0
```

```
Exit Do  
Loop  
End If
```

```
If k4 >= 65536 Then  
Do While k5 < 65536
```

```
s2 = "c" & k5 + 1  
s1 = "a" & k5 + 1  
Sheets("transit").Select  
Range("aq5").Select  
Selection.Copy  
Sheets("Calculine5").Select  
Range(s2).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False
```

```
Sheets("transit").Select  
Range("a3").Select  
Selection.Copy  
Sheets("calculine5").Select  
Range(s1).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False  
k5 = k5 + 1  
k = 0
```

```
Exit Do  
Loop  
End If
```

```
Next II
```

```
Sheets("board").Select
```

→ This sub routine allows to recover number of compared peaks in "AM" column in "board" sheet for each comparison. Calculations was done in "CommandButton1_Click" sub macro by clicking on Alt2 button.

```
Range(Cells(ii, 39), Cells(20, 39)).Select  
Selection.Copy  
Sheets("calculine").Select  
Cells(jj, 5).Select  
Selection.PasteSpecial Paste:=xlValues, Operation:=xlNone, SkipBlanks:= _  
False, Transpose:=False  
jj = jj + nbsamp  
ii = ii + 1
```

```
nbsamp = nbsamp - 1
```

ANNEXE 1


l1 = i + 3


Next i


End Sub

If you have any questions or suggestions, please don't hesitate to contact me at :

Dr. Laurence Dujourdy
Institut de Police Scientifique et de Criminologie
UNIL - Bâtiment de Chimie
BCH, 1015 Lausanne - Dorigny
Switzerland

 (+ 41 21) 692 46 49

 (+ 41 21) 692 46 05

 Laurence.Dujourdy@ipsc.unil.ch

ANNEXE 2

Page de garde et sommaire du projet CHAMP



CIS8-CT-2004-502126

CHAMP

Collaborative Harmonisation of Methods for profiling of Amphetamine Type Stimulants

Instrument: Specific Targeted Project
Thematic Priority: 8.1 Policy-oriented research

FINAL SCIENTIFIC REPORT

Collaborative Harmonisation of Methods for profiling of Amphetamine Type Stimulants

Actual Submission date: 12.3.2007

Start date of project: 1.7.2004

Duration: 1.7.2004-30.9.2006

Organisation name of lead contractor for this deliverable: NBI

Revision [1]

Project co-funded by the European commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	X

TABLE OF CONTENTS

1	INFORMATION ON THE ILLICIT AMPHETAMINE TYPE STIMULANTS (ATS) TRAFFIC AND PRODUCTION IN PARTNER'S AND OTHER ATS PRODUCING COUNTRIES	6
1.1	INTRODUCTION	6
1.2	TOPICS AND TIME PERIOD OF TRACING ATS SITUATION	6
1.3	SURVEY OF RECEIVED DATA	7
1.3.1	<i>European countries</i>	7
1.3.2	<i>The U.S.A.</i>	11
1.3.3	<i>Australia</i>	12
1.4	DISCUSSION	15
1.5	CONCLUSIONS	16
2	GUIDELINES FOR MEASUREMENT OF EXTERNAL FEATURES	17
2.1	INTRODUCTION	17
2.2	GENERAL INFORMATION	17
2.3	THE STRUCTURE OF THE DATA SHEET	18
2.3.1	<i>Visual description</i>	19
2.3.2	<i>Measurement of dimensions</i>	20
2.4	IMAGE INFORMATION	20
3	IDENTIFICATION OF THE ACTIVE SUBSTANCE AND THE ORGANIC ADULTERANTS AND DILUENTS PRESENT IN ATS SAMPLES	21
3.1	INTRODUCTION	21
3.2	RESULTS	22
3.3	RECOMMENDATIONS	27
3.3.1	<i>Organic adulterants</i>	27
3.3.2	<i>Sugars</i>	27
3.3.3	<i>Fatty Acids</i>	28
3.3.4	<i>Other diluents</i>	28
3.4	ANALYSIS OF MDMA STREET SAMPLES	29
3.4.1	<i>Organic adulterants:</i>	29
3.4.2	<i>Organic diluents</i>	30
4	IDENTIFICATION OF THE ELEMENTS PRESENT IN ATS SAMPLES	33
4.1	GENERAL INTRODUCTION	33
5	POLARISED XRF	34
5.1	ELEMENTAL ANALYSIS BY ICP-MS/AES (TASK 2.4.1)	36
5.1.1	<i>Introduction to ICP-MS and ICP-AES analysis</i>	36
5.1.2	<i>Description and comparison of the ICP-MS procedures used by NFI and DEA (T2.4.1.1)</i>	37
5.1.3	<i>Analysis of control samples (T2.4.1.2)</i>	37
5.1.4	<i>Analysis of MDMA and MA test samples, evaluation of inorganic profiling (T2.4.1.3)</i>	38
5.1.5	<i>Measurement of the 80 MDMA street samples (T2.4.1.4)</i>	41
5.1.6	<i>Conclusions for the analysis of MDMA and MA samples by ICP-MS/AES</i>	42
5.1.6.1	<i>Comparison of Turboquant-XRF with ICP-MS and ICP-AES results (T2.4.1.5)</i>	42
5.2	ELEMENTAL AND ION ANALYSIS BY XRF, CE AND WET CHEMISTRY (TASK 2.4.2)	44
5.2.1	<i>Introduction to XRF analysis</i>	44
5.2.2	<i>Comparison and adaptation of different XRF systems in partners' laboratories (T2.4.2.1)</i>	45
5.2.3	<i>Identification of salt form in MDMA and MA control, test and street samples (T2.4.2.2)</i>	47
5.2.4	<i>Analysis of MDMA and MA control, test and street samples with XRF (T2.4.2.3)</i>	49
5.2.5	<i>Additional analysis of MDMA and MA control and test samples with CE (T2.4.2.4)</i>	50
5.2.6	<i>Correlation of XRF and CE elemental data with FT-IR and XRD data (T2.4.2.5)</i>	52
5.2.7	<i>Development of a harmonised polarised-XRF procedure (via Turboquant) and application to MDMA, MA control, test and street samples (T2.4.2.6)</i>	54
5.3	FINAL CONCLUSIONS AND RECOMMENDATIONS FROM WP2.4:	55
6	QUANTITATION OF THE ACTIVE SUBSTANCE IN ATS SAMPLES	56

6.1	INTRODUCTION	56
6.2	SAMPLE PREPARATION	56
6.3	QUANTITATIVE STANDARDS	57
6.4	QUANTIFICATION METHODS AND TECHNIQUES	57
6.5	RESULTS	58
6.6	INTER-CALIBRATION EXERCISE	59
6.7	ANALYSIS OF STREET SAMPLES	62
6.8	CONCLUSIONS	62
7	DETERMINATION OF ENANTIOMERIC COMPOSITION OF ATS SAMPLES	63
7.1	INTRODUCTION	63
7.2	AGREEMENT ON PREPARATIVE METHOD AND MEASUREMENT METHOD	63
7.3	REPORTING METHOD	63
7.4	ANALYSIS OF 20 TEST SAMPLES	63
7.5	COMPARISON OF THE DATA	66
7.6	FINAL METHOD ADJUSTMENTS	67
7.7	MEASUREMENT OF THE 80 SAMPLES	67
8	OPTIMISATION OF GC-MS PROFILING METHOD FOR MDMA	68
8.1	INTRODUCTION	68
8.2	ACCOMPLISHED TASKS	68
8.2.1	<i>Circulation of methods used in various labs + work program + proposal for a method (T3.2.1.1)</i>	68
8.2.2	<i>Testing of instrumentation with a modified Grob sample and preliminary testing with 2 MDMA control samples. Optimisation of sample preparation procedures (NBI and NFI). Testing with the amphetamine profiling method (T 3.2.1.2).</i>	69
8.2.2.1	Grob sample	69
8.2.2.2	MDMA control samples 1 and 2	70
8.2.2.3	Filtering tests	72
8.2.2.4	Testing with the amphetamine profiling method	75
8.2.3	<i>Training workshop for partners participating in the task (T 3.2.1.3)</i>	81
8.2.4	<i>Validation of method by all participating labs. Analyses of four MDMA control samples. Evaluation of repeatability and reproducibility within and between labs (T3.2.1.4)</i>	81
8.2.5	<i>Analysis of the 26 MDMA test samples. Integration and evaluation of results (T3.2.1.5)</i>	83
8.2.6	<i>Testing of instrumentation with a modified Grob sample and preliminary testing with 2 MDMA control samples. Optimisation of sample preparation procedures (NBI and NFI). Testing with the amphetamine profiling method (T3.2.1.6)</i>	85
8.2.6.1	Stability studies	85
8.2.6.2	Measurements on a 27 th test sample	86
8.2.6.3	Integration exercise	87
8.2.6.4	Tentative identification of impurities	89
8.3	SUMMARY, OVERALL DISCUSSIONS AND CONCLUSIONS	95
9	OPTIMISATION OF GC-MS PROFILING METHOD FOR METHAMPHETAMINE	97
9.1	INTRODUCTION	97
9.2	ACCOMPLISHED TASKS	97
9.2.1	<i>Determination of potential target compounds and preliminary testing (T 3.2.2.1)</i>	97
9.2.1.1	Determination of potential target compounds	98
9.2.1.2	Preliminary testing	100
9.2.2	<i>Circulation of method and work programme (T3.2.2.2)</i>	106
9.2.3	<i>Training workshop (T.3.2.2.3)</i>	106
9.2.4	<i>Validation of method including evaluation of variability within and between labs (T.3.2.2.4)</i>	106
9.2.4.1	Grob sample	106
9.2.4.2	MA control samples	108
9.2.4.3	MA test samples	113
9.2.4.4	MA Street samples	113
10	OPTIMISATION OF IRMS METHOD(S) FOR MDMA & METHAMPHETAMINE SAMPLES	114
10.1	INTRODUCTION	114
10.2	EXPERIMENTAL SECTION	115
10.2.1	<i>Sample Preparation</i>	115
10.2.2	<i>Stable isotope analysis of elements C, N, H at BKA</i>	115
10.2.3	<i>Experimental conditions used for GC-IRMS at IHCP</i>	116
10.3	EVALUATION OF RESULTS	116

10.3.1	<i>EA-IRMS</i>	116
10.3.2	<i>GC-IRMS analyses</i>	121
11	MULTIVARIATE STATISTICAL ANALYSIS	124
11.1	INTRODUCTION	124
11.1.1	<i>Objectives of Task 4.2</i>	124
11.1.2	<i>Subtasks of Task 4.2</i>	124
12	EXTERNAL CHARACTERISTICS AND CHEMICAL COMPOSITION OF MDMA TABLETS (WP2)	127
12.1	SELECTION OF VARIABLES	127
12.1.1	<i>Reliability of variables</i>	127
12.1.1.1	External characteristics of MDMA tablets (T2.2)	127
12.1.1.2	Organic composition (T2.3)	129
12.1.1.3	Inorganic composition (T2.4)	129
12.1.1.4	Quantitative Analysis (purity in percent) (T2.5)	129
12.1.1.5	Enantiomeres (T2.6)	130
12.1.1.6	Summary	131
12.1.2	<i>Relevance of variables</i>	131
12.1.3	<i>Selected variables</i>	134
12.2	DISCRIMINATING POWER OF VARIABLES	134
12.2.1	<i>Correlation coefficients</i>	134
12.2.2	<i>Discriminant analysis</i>	134
12.2.3	<i>Statistical pre-treatments</i>	136
12.2.4	<i>Cluster analysis</i>	137
12.3	DISTRIBUTIONS OF LINKED AND UNLINKED SAMPLES (TEST SAMPLES)	138
12.3.1	<i>ROC Curves</i>	138
12.3.1.1	Within laboratory distributions	138
12.3.1.2	Between laboratories distributions	140
12.3.1.3	Comparison of within and between laboratories distributions	141
12.3.2	<i>Histograms</i>	142
12.4	DISTRIBUTIONS OF LINKED AND UNLINKED SAMPLES (STREET SAMPLES)	143
12.4.1	<i>ROC Curves</i>	143
12.4.2	<i>Histograms</i>	144
12.5	CLANDESTINE TRAFFIC PATTERNS	147
12.5.1	<i>Descriptive statistics</i>	147
12.5.2	<i>Discriminant analysis</i>	148
12.5.3	<i>Principal Component Analysis (PCA)</i>	148
12.6	CONCLUSION	150
13	IMPURITY PROFILES OF MDMA TABLETS (WP3)	151
13.1	SELECTION OF TARGET COMPOUNDS	151
13.2	DISCRIMINATING POWER OF VARIABLES	153
13.2.1	<i>Discriminant analysis</i>	153
13.2.2	<i>Principal Component Analysis (PCA)</i>	155
13.3	DISTRIBUTIONS OF LINKED AND UNLINKED SAMPLES	156
13.3.1	<i>ROC Curves</i>	156
13.3.1.1	Between laboratories distributions	156
13.3.1.2	Within laboratory distributions	157
13.3.2	<i>Histograms</i>	158
13.4	CLANDESTINE TRAFFIC PATTERNS	161
13.5	CONCLUSION	163
14	DRUG INTELLIGENCE	164
14.1	ORGANIC IMPURITY PROFILING	165
14.2	EXTERNAL CHARACTERISTICS PROFILING	168
14.3	CONCLUSION	169
15	IMPURITY PROFILES OF METAMPHETAMINE (WP3)	170
15.1	SELECTION OF TARGET COMPOUNDS	170
15.1.1	<i>Stability study</i>	170
15.1.2	<i>Study of correlation between target compounds</i>	176
15.1.3	<i>Final selection of target compounds combining correlation, stability and integration criteria</i>	176

15.2	DISTRIBUTIONS OF LINKED AND UNLINKED SAMPLES	181
15.2.1	<i>Study of correlations (Pearson and Cosine) with new data (24 target compounds)</i>	181
15.3	ROC CURVES	189
15.3.1	<i>Histograms</i>	189
15.4	SYNTHETIC ROUTE PATTERNS	192
15.4.1	<i>PCA analysis on 24 target compounds data</i>	192
15.4.2	<i>PCA analysis on Street Samples</i>	197
15.4.3	<i>Cluster analysis on Street Samples made from the ephedrine synthesis route</i>	199
15.4.4	<i>PLS-DA analysis on 24 target compounds data</i>	203
15.5	CONCLUSION	209
	ANNEX I: OVERVIEW MDMA PROFILING PROCEDURES	210
	ANNEX II: MDMA PROFILING METHOD	212
	ANNEX III: RSDS FOR MDMA IMPURITIES	215
	ANNEX IV: JRC'S REPORT ON THE ISOTOPIC DETERMINATION ON ILLICIT DRUGS	217
	ANNEX V: RELEVANT LITERATURE FOR THE USE OF ICP-MS AND/OR ICP-AES AND XRF, EDXRF, TXRF AND XRD	226
	ANNEX VI: DIGESTION PROCEDURES FOR ICP-MS MEASUREMENT BY DEA AND NFI	228
	ANNEX VII: INTER-LAB VARIATION IN THE ELEMENT CONCENTRATIONS IN THE MDMA AND METHAMPHETAMINE TEST SAMPLES	230
	ANNEX VIII: PT CONCENTRATIONS IN TEST SAMPLES BY ICP-MS (NFI).	232
	ANNEX IX: COMPARISON OF XRF WITH ICP-AES / ICP-MS (AND CE⁺)	233
	ANNEX X: ED-XRF SYSTEMS AND INSTRUMENT PARAMETERS USED BY BKA, NFI, ICP INSTRUMENT	234
	INSTRUMENT	234
	ANNEX XI: SCHEME FOR THE IDENTIFICATION OF ANIONS VIA WET CHEMISTRY (PROPOSAL BY BKA)	237
	ANNEX XII: COMPARISON OF XEPOS (BKA) AND X-LAB (NFI) ELEMENTAL DATA	238
	ANNEX XIII: PROCEDURES FOR CE ANION AND CATION ANALYSIS OF ATS BY BKA	241
	ANNEX XIV: CORRELATION OF XRF AND CE DATA WITH FT-IR AND XRD DATA FOR ALL MDMA CONTROL AND TEST SAMPLES AND SOME MA SAMPLES	242
	ANNEX XV: STANDARD XRF-XEPOS ELEMENTAL DATA FOR BKA'S MDMA STREET SAMPLES	243
	ANNEX XVI: STANDARD XRF-X-LAB ELEMENTAL DATA FOR NFI'S MDMA STREET SAMPLES	244
	ANNEX XVII: TURBOQUANT-XRF DATA FOR MDMA AND MA C+T SAMPLES (BKA, NFI)	245

ANNEXE 3

Page de garde et sommaire du rapport final pour le
Fonds National Suisse pour la Science (FNS)



SWISS NATIONAL SCIENCE FOUNDATION

**ECOLE DES SCIENCES CRIMINELLES
INSTITUT DE POLICE SCIENTIFIQUE**

FINAL REPORT

Responsible applicant	Dr DUJOURDY Laurence
Co-applicants	Pr. TARONI Franco – Pr. MARGOT Pierre
Project title	Classification, interpretation and statistical evaluation of data in illicit drug profiling
Request number	1115-068017/2
Period	01.11.2002 - 31.10.2003

Abstract

In the identification process, there is an important step useful in drug intelligence, that is: the determination of some set to which an object/sample belongs, as the determination and to whether an object/sample belongs to this given set. How the evidence is presented to a Court of Law, and how links between seizures can be detected and represented are specific challenges in forensic science.

Most methods used to determine traffic organisations, drug rings etc. are based on traditional police investigations (observation, tapping, interrogation). Evidence of organised crime is usually difficult to establish and contested. One way to overcome these drawbacks is to highlight extended activity linked to a common or unique source of products or treatments that will connect multiple seizures of samples within an organised traffic. The analysis of chemical profile is seen to be an objective means of demonstrating such links.

The objectives of this research were to develop a strategy to control the presence of physico-chemical links between seizures of heroin, cocaine, amphetamines and derivatives. Such work has already been dedicated to this problem especially concerning the analytical optimisation of parameters and impurities detection observed in a drug seizure but little research has addressed the more difficult problem dealing with, (1), the data treatment using different comparisons and pattern recognition methods and (2), management, visualisation with the intention to deliver a chemical link information in a strategic context.

In general, this study looked into the interpretation of link evidence between seizures.

This report is divided as follow:

- The two first part are an introduction followed by a rapid current state of the art description,
- The third part concerns the research carried out by the applicant which is divided in several sections :
 - analytical with the optimization of Gas Chromatograph parameters for the comparison between two machines,
 - statistical with the use of chemometrics for the data analysis,
 - visualisation and management of links.

Key words: organised crime, illicit drugs, gas chromatography, drug profiling, drugs intelligence, pattern recognition, classification, interpretation of evidence.

Table of contents

I. INTRODUCTION	3
II. CURRENT STATE OF THE ART	4
II.1. ILLICIT DRUG PROFILING	4
II.2. DATA ANALYSIS	4
II.3. INTERPRETATION OF DATA.....	5
II.4. USE OF THE INFORMATION IN THE INVESTIGATION	5
III. STATE OF THE RESEARCH CARRIED OUT BY APPLICANTS	6
III.1. OPTIMIZATION OF GAS CHROMATOGRAPH PARAMETERS FOR THE COMPARISON OF ANALYSES BETWEEN TWO APPARATUS.....	6
III.1.1. Analytical methods used with reference Autosystem model for heroin and cocaine impurities separation.....	6
III.1.2. Analytical methods used with reference Autosystem XL model for heroin and cocaine impurities separation	7
III.2. DATA TREATMENT AND CHEMOMETRICS.....	20
III.2.1. Heroin samples	20
III.2.2. Cocaine samples	38
III.2.3. Amphetamines and derivates.....	52
III.3. INTERPRETATION OF DATA.....	55
III.4. USE OF THE INFORMATION IN THE INVESTIGATION.....	59
IV. CONCLUSION	68
BIBLIOGRAPHY	70