



UNIVERSITE SULTAN MOULAY SLIMANE
Faculté des Sciences et Techniques
Béni-Mellal



Centre d'Études Doctorales : Sciences et Techniques

Formation Doctorale : Mathématiques et Physiques Appliquées

THÈSE

Présentée par

AMKOR ALI

Pour l'obtention du grade de

DOCTEUR

Discipline : Physique

Spécialité : Génie Electrique

Conception et réalisation d'un système multi-capteur pour préserver la santé humaine des dangers des aliments traités

Soutenue le Vendredi 03 Juin 2022 à 10h devant la commission d'examen :

Pr. Mohammed SAJIEDDINE	Professeur, Université Sultan Moulay Slimane, E.N.S.A Khouribga, Maroc	Président
Pr. Cherki DAOUI	Professeur, Université Sultan Moulay Slimane, F.S.T Béni-Mellal, Maroc	Rapporteur
Pr. Abdellah HALIMI	Professeur, Université Hassan Premier I.S.S.S. Settat, Maroc.	Rapporteur
Pr. Abderrazak LFAKIR	Professeur, Université Sultan Moulay Slimane, F.S.T Béni-Mellal, Maroc	Rapporteur
Pr. Karim RHOFIR	Professeur, Université Sultan Moulay Slimane, E.N.S.A Khouribga, Maroc.	Examineur
Pr. Nouredine EL BARBRI	Professeur, Université Sultan Moulay Slimane, E.N.S.A Khouribga, Maroc	Directeur de Thèse

A

- *Mon adorable Maman*
- *Ma chère femme Assiya et mon cher fils Ibrahim Khalil*
- *Mon cher Oncle Driss, ma grande mère, ma petite sœur Hiba et
toute ma famille*
 - *Mes amis*
- *Mes collègues (Administratifs et Professeurs) à l'ENSA de
Khouribga*
- *Mes frères les agents de sécurité, les jardiniers et les femmes de
ménages de l'école ENSA Khouribga*
 - *À tous ceux qui ont fait de moi ce que je suis*

REMERCIEMENTS

Tout d'abord, je remercie mon Dieu le Tout-Puissant pour toutes ses bénédictions et de m'avoir donné la force pour terminer ce travail de thèse.

Après plus de 4 ans de travail acharné et continu, il est enfin temps d'adresser quelques mots de gratitude aux personnes sans qui ce travail n'aurait peut-être pas été possible puisqu'il ne remercie pas Dieu qui ne remercie pas les gens.

Tout d'abord, je voudrais adresser ma plus profonde reconnaissance à mon Directeur de thèse Monsieur **Noureddine El Barbri**, Professeur à l'Ecole Nationale des Sciences Appliquées de Khouribga relevant de l'Université Sultan Moulay Slimane de Beni Mellal, pour sa confiance en moi, ses conseils, son soutien, ses encouragements et sa disponibilité pour la réussite de notre travail. Professeur **Noureddine El Barbri** m'a donné l'opportunité de travail dans le domaine de recherche à ses côtés et il a préparé un environnement de travail passionnant avec de nombreuses opportunités pour explorer de nouvelles idées. Cela n'aurait pas été possible sans ses orientations et conseils professionnels, je le remercie encore pour sa manière efficace et efficiente de direction et j'espère continuer à collaborer avec ce grand homme dans d'autres problématiques de recherche à venir, si Dieu le veut.

Je remercie les membres de mon jury d'avoir offert leur temps et d'avoir accepté de juger mon travail.

Je tiens à remercier tout le corps professoral du département Génie Electrique de l'Ecole Nationale des Sciences Appliquées de Khouribga pour leurs aides et leurs prestigieux conseils en particulier Professeur **Karim Rhofir** qui n'a pas cessé de me prodiguer ses conseils et me transmettre son énergie positive avec son habituelle gentillesse à chaque fois que le besoin se ressentait.

Mes remerciements vont également au corps administratif, mes collègues de l'Ecole Nationale des Sciences Appliquées de Khouribga, où j'ai commencé à travailler il y a plus de dix ans, et grâce à leur soutien et leurs conseils, j'ai pu, Dieu merci, progresser jusqu'à l'obtention du diplôme de doctorat.

Enfin, un grand merci à ma petite famille, en particulier ma mère, ma femme et mon oncle Idriss pour leur amour, patience et leur foi en moi, à ma grande famille et à tous mes amis.

Résumé

L'identification des produits alimentaires, pollués et contaminés par des produits chimiques, ayant un impact négatif sur la santé humaine, est importante pour les consommateurs. Les travaux de recherches que nous rapportons dans cette thèse s'inscrivent dans le cadre du développement d'un dispositif « nez électronique portable », miniaturisé et à moindre coût, basé sur des capteurs de gaz orienté vers des applications visant à vérifier la qualité des produits alimentaires, notamment les produits traités avec des pesticides. Dernièrement, les nez électroniques ont été largement exploités dans différents domaines y compris l'agroalimentaire grâce à leur efficacité, leur précision, leur moindre coût et leur facilité d'utilisation, contrairement aux méthodes physico-chimiques. Les nez électroniques sont constitués de deux parties principales, une partie matérielle qui contient principalement un réseau de capteurs de gaz et une partie logicielle dédiée au traitement et classification de données.

Après une étude de marché sur les capteurs de gaz, notre choix s'est porté sur des capteurs de gaz à oxyde métallique du commerce qui seront utilisés pour la réalisation de notre dispositif. Dans un premier temps, une méthodologie de conception et d'impression de deux circuits est présentée, qui vise à optimiser le dimensionnement et l'agencement des capteurs d'un côté et à l'adaptation et la protection des composants d'un autre côté. Nous avons établi un protocole de mesure en se fixant des contraintes de sensibilité et de réponse. Un prototype avec une chambre en verre pour l'isolation, conçu d'après cette méthodologie, est testé, réalisé, mis en œuvre et caractérisé.

L'expérimentation de notre dispositif pour des échantillons de menthe traitée avec deux dangereux insecticides nous a permis de distinguer entre les échantillons avec des taux de réussites allant de 75% à 100% en utilisant la méthode des Séparateurs à Vaste Marges (SVM) dans un premier temps et K plus Proches Voisins (KNN) dans un autre temps et une prédiction de jour de traitement pour un insecticide avec un taux de corrélation qui est arrivé à 99% grâce à la méthode de Réseau Autorégressif non linéaire à entrées exogènes (NARX).

Ensuite, nous nous sommes intéressés à discriminer entre trois échantillons de pommes de terre selon leurs champs de culture : traité traditionnellement avec le fumier de moutons et d'ânes, traité avec des engrais NPK et traité avec le fumier de poulet. En appliquant la méthode SVM, nous avons atteint un taux de réussite qui dépasse 92%.

Au final, on peut dire que les nez électroniques ainsi développés se sont révélés être des outils de surveillance fiables de la contamination des produits alimentaires.

Mots clés : nez électronique, capteurs de gaz à oxyde métallique, discrimination entre la menthe traitée et non traitée, distinction des pommes de terre par le champ de culture, analyse de données, méthodes de reconnaissances de formes.

Abstract

The identification of food products polluted and contaminated by chemicals with a negative impact on human health is important for consumers. The research we report in this thesis is part of the development of a portable, miniaturized, low-cost electronic nose device, based on gas sensors oriented to applications aimed at verifying the quality of food products, especially the ones treated with pesticides. Recently, electronic noses have been widely exploited in various fields including the food industry thanks to their efficiency, precision, and ease of use, unlike physicochemical methods. The electronic noses consist of two main parts, a hardware one which mainly contains a network of gas sensors, and the second consists of a data processing and classification.

After a benchmarking on gas sensors, our choice fell on commercial metal oxide gas sensors which will be used for the realization of our device. First, a methodology for the design and printing of two circuits is presented, which aims to optimize the sizing and arrangement of the sensors on one side and the adaptation and protection of the components on the other side. We have established a measurement protocol by setting sensitivity and response constraints. A prototype with a glass chamber for insulation, designed according to this methodology, is tested, produced, implemented, and characterized.

The experimentation of our device for samples of mint treated with two dangerous insecticides allowed us to discriminate between the samples with success rates ranging from 75% to 100% using the Support Vector Machines (SVM) method at first and the K Nearest Neighbors (KNN) method at another time and the treatment day prediction for an insecticide with a correlation rate that arrived at 99% thanks to the NARX method.

Then, we were interested in discriminating between three potato samples according to their crop fields: traditionally treated with sheep and donkey manure, treated with NPK fertilizers, and treated with chicken manure. By applying the Support Vector Machines (SVM), we achieved a success rate that exceeds 92%.

In the end, it can be said that the electronic noses developed in this way have proved to be reliable monitoring tools for monitoring food products' contamination.

Keywords: electronic nose, metal oxide gas sensors, discrimination between treated and untreated mint, the distinction of potatoes by field of cultivation, data analysis, shape recognition methods.

Liste des figures

Figure 1-1 Système olfactif humain ; Illustration de Patrick J. Lynch, illustrateur médical....	23
Figure 1-2 Comparaison schématique de la technologie du nez électronique et du système olfactif [23].....	28
Figure 1-3 Architecture d'un capteur de gaz à oxyde métallique de la société Figaro, (a) la structure la plus ancienne [29] et (b) la structure récente [30].....	31
Figure 1-4 Influence de la variation de la température sur la réponse du capteur MQ-7 de commerce à oxyde métallique SnO ₂	35
Figure 1-5 Influence de la variation de l'humidité et la température sur la réponse du capteur TGS 2611E00 du commerce	36
Figure 1-6 Schéma de principe d'un chromatographe en phase gazeux.....	41
Figure 1-7 Principe d'un spectromètre de masse [59].....	42
Figure 1-8 Chromatographe en phase gazeuse couplée à un spectromètre	43
Figure 1-9 Chromatographie en phase gazeuse couplée à l'Olfactométrie [60]	43
Figure 2-1 Exemple d'un hyperplan optimal dans \mathcal{R}^2	59
Figure 2-2 Représentation de l'astuce du changement de l'espace	60
Figure 2-3 Problème du choix de facteur k pour la méthode KNN	63
Figure 2-4 Matrice de confusion et les mesures conventionnelles.....	65
Figure 2-5 Conception d'un neurone formel d'après McCulloch et Pitts	68
Figure 2-6 Exemple d'un réseau multicouche	70
Figure 2-7 Réseau de NARX à deux couches	71
Figure 2-8 Structure du réseau NARX.....	72
Figure 2-9 Architectures possibles du réseau NARX, avec U(t) valeur de l'entrée, y(t) valeur de sortie, $\hat{y}(t)$ la valeur de sortie estimée et TDL est la ligne à retard à prises.....	72
Figure 3-1 Conception ouverte d'un nez électronique.....	82
Figure 3-2 Conception statique d'un nez électronique	83
Figure 3-3 Conception dynamique d'un nez électronique.....	84
Figure 3-4 Exemple des différents capteurs présents dans les marchés.....	86
Figure 3-5 Matrice de capteurs réalisée	89
Figure 3-6 Montages électroniques correspondant aux types de capteurs choisis, (a) pour les capteurs à 2 broches, (b) pour les capteurs à 4 broches et (c) pour les capteurs à 6 broches... ..	89
Figure 3-7 Carte d'acquisition de données ADLINK USB-1901	92
Figure 3-8 Montage d'adaptation et de protection pour chaque capteur	93
Figure 3-9 Carte d'adaptation et de protection réalisée	93
Figure 3-10 Schéma fonctionnel du programme LabVIEW réalisé et utilisé pour l'acquisition de données	94
Figure 3-11 Face-avant du programme LabVIEW d'acquisition de données	95
Figure 3-12 Représentation schématique de l'outil utilisé	95
Figure 4-1 Réponse des capteurs en présence de différents types de menthe.....	104
Figure 4-2 Résultats de l'ACP des données de capteurs obtenues à partir de : (a) ensemble de données X1 et X2, (b) jeux de données X1 et X3, (c) ensembles de données X2 et X3 et (d) ensembles de données globales X4	107

Figure 4-3 Résultats de la classification SVM des différents cas de menthe (a) menthe non traitée contre menthe traitée par Decis Fluxx, (b) menthe non traitée contre menthe traitée par Malyphos, (c) menthe traitée par Decis contre menthe traitée par Malyphos.....	109
Figure 4-4 Matrice de confusion pour différents cas de menthe, (a) menthe non traitée contre menthe traitée par Decis Fluxx, (b) menthe non traitée contre menthe traitée par Malyphos, (c) menthe traitée par Decis contre menthe traitée par Malyphos, (d) menthe non traitée contre menthe traitée	110
Figure 4-5 Réponses temporelles des trois capteurs de gaz pour la menthe non traitée et la menthe traitée en fonction des jours.....	114
Figure 4-6 Projection des données sur les trois premières composantes principales de l'ACP	116
Figure 4-7 Résultat de KNN pour l'identification du type de menthe.....	118
Figure 4-8 Taux de réussite de KNN pour l'identification du type de menthe.....	118
Figure 4-9 Architecture du réseau NARX dans Matlab	120
Figure 4-10 Performance du réseau NARX	121
Figure 4-11 Corrélation entre les sorties et les valeurs cibles. (a) pour les données d'apprentissage, (b) pour le test et (c) pour toutes les données	122
Figure 4-12 Résultat de corrélation de l'algorithme PLSR pour la prédiction du jour du prélèvement de la menthe traitée.....	123
Figure 4-13 Résultat de corrélation de l'algorithme ϵ -SVM pour la prédiction du jour du prélèvement de la menthe traitée.....	124
Figure 5-1 Exemples des échantillons préparés et utilisés	132
Figure 5-2 Exemples des réponses enregistrées par les cinq capteurs à 5 échantillons de pommes de terre prélevés sur le champ traité traditionnellement et à 5 autres échantillons prélevés sur le champ traité par les engrais NPK.....	133
Figure 5-3 Nuage de points PCA des échantillons de pommes de terre selon leurs sources, traités traditionnellement ou traités avec des engrais NPK.....	135
Figure 5-4 Résultat de la classification SVM pour les échantillons de pommes de terre selon leurs sources, non traitées ou traitées avec des engrais NPK.....	136
Figure 5-5 Analyses microbiologiques des échantillons de pommes de terre	139
Figure 5-6 Exemples des échantillons préparés et utilisés	141
Figure 5-7 Exemples de réponses enregistrées par les cinq capteurs à cinq échantillons de pommes de terre de chaque type, la couleur verte représente celles prélevées du champ traité avec des engrais traditionnels, la couleur orange pour celles traitées avec des engrais NPK et la couleur bleue pour celles traitées avec du fumier de poulet.....	142
Figure 5-8 Résultats SVM pour l'identification des types de pommes de terre par la première approche, (a) représentation graphique du résultat et (b) matrice de confusion	145
Figure 5-9 Résultat de la première étape de la deuxième approche, (a) représentation du résultat de la discrimination des pommes de terre traitées traditionnellement de celles traitées différemment et (b) la matrice de confusion	147
Figure 5-10 Résultat de la deuxième étape de la deuxième approche, (c) représentation du résultat de la discrimination des pommes de terre traitées avec du fumier de poulet et de celles traitées avec des engrais NPK et (b) la matrice de confusion	148

Liste des tableaux

Tableau 1-1 Différents types des capteurs, leurs matériaux sensibles et leurs principes de fonctionnement.....	29
Tableau 1-2 Mécanisme de réaction dans les capteurs de gaz à oxyde métallique.....	32
Tableau 1-3 Nez électroniques commercialisés et leurs domaines d'utilisation [56].....	38
Tableau 2-1 Caractéristiques les plus utilisées dans l'extraction des données pertinentes et leurs descriptions.....	53
Tableau 2-2 Méthodes les plus utilisées dans la normalisation des données.....	54
Tableau 3-1 Capteurs utilisés, leurs gaz cibles et prix.....	87
Tableau 4-1 Résultats de l'étude comparative des méthodes de régression.....	125

Table des matières

INTRODUCTION GENERALE.....	15
Chapitre 1 : METHODES D'ANALYSE DES ODEURS ET CAPTEURS DE GAZ.....	18
I. INTRODUCTION.....	19
II. LES ODEURS.....	20
II.1. Définition.....	20
II.2. Rôle de l'odeur	20
II.3. Propriétés des odeurs	21
II.4. Paramètres des odeurs	21
III. LE NEZ HUMAIN.....	22
III.1. Architecture du système olfactif humain	22
III.1.1. L'épithélium olfactif	23
III.1.2. Bulbe olfactif.....	24
III.1.3. Cortex olfactif	24
III.2. Facteurs influençant la perception des odeurs.....	25
III.3. Limite d'utilisation du nez humain	25
IV. LE NEZ ELECTRONIQUE :.....	26
IV.1. Définition	26
IV.2. L'historique du nez électronique :	26
IV.3. Principe de fonctionnement du nez électronique	27
IV.4. Capteurs de gaz	28
IV.4.1. Les différents types de capteurs de gaz.....	28
IV.4.2. Capteurs résistifs à oxydes métalliques semi-conducteurs (MOX)	29
IV.4.2.1. Historique.....	30
IV.4.2.2. Structure d'un capteur à oxydes métalliques	30
IV.4.2.3. Principe de fonctionnement	31
IV.4.3. Caractéristiques d'un capteur de gaz à oxyde métallique (MOX).....	33
IV.4.3.1. Sensibilité.....	33
IV.4.3.2. Sélectivité.....	33
IV.4.3.3. Stabilité	34
IV.4.4. Facteurs affectant les capteurs de gaz à oxyde métallique (MOX).....	34
IV.4.4.1. Température	35
IV.4.4.2. Humidité.....	35

IV.4.5. Critères de sélection d'un capteur de gaz à oxyde métallique.....	36
IV. 4.6. Champs d'applications	37
IV. 4.7. Quelques nez électroniques commercialisés.....	38
V. QUELQUES METHODES ANALYTIQUES D'ANALYSE DES ODEURS	40
V.1. Chromatographie en phase gazeuse.....	40
V.2. Spectrométrie de masse	41
V.3. Chromatographie en phase gazeuse couplée à la spectrométrie de masse	42
V.4. Chromatographie en phase gazeuse couplée à l'Olfactométrie.....	43
VI. CONCLUSION	44
Références	44
Chapitre 2 : ANALYSE DE DONNEES ET APPRENTISSAGE AUTOMATIQUE	50
I. INTRODUCTION.....	51
II. PRETRAITEMENT DE DONNEES	52
II.1. Le centrage et la mise à l'échelle	52
II.2. Extraction des caractéristiques (Features extraction)	52
II.3. Normalisation	54
III. METHODES D'ANALYSES MULTIVARIABLES DE DONNEES	55
III.1. Méthodes de classification des données.....	55
III.1.1. Analyse en Composantes Principales (ACP)	55
III.1.1.1. Principe de l'ACP.....	56
III.1.1.2. Analyse du nuage des points-individus et points variables.....	57
III.1.1.3. Algorithme de l'ACP	58
III.1.2. Séparateurs à Vaste Marge (SVM).....	58
III.1.2.1. SVM avec classificateur linéaire.....	59
III.1.2.2. SVM avec classificateur non-linéaire.....	60
III.1.2.3. Classification multi-classes par SVM	60
III.1.2.4. Algorithme de SVM	62
III.1.3. K plus Proches Voisins (KNN)	62
III.1.3.1. Présentation et principe de la méthode KNN	62
III.1.3.2. Algorithme de KNN	64
III.2. Indicateurs d'évaluation des méthodes de classification.....	65
III.3. Méthodes de régression des données	65
III.3.1. Régression par moindres carrés partielles (PLSR).....	65
III.3.2. Régression linéaire des Séparateurs à Vaste Marge (ϵ -SVM).....	67
III.3.3. Réseau Autorégressif non linéaire à entrées exogènes (NARX).....	68
III.3.3.1. Généralités sur les réseaux de neurones artificielles (RNA).....	68
III.3.3.2. Perceptron Multicouche (PMC)	69

III.3.3.3. Réseau NARX	71
III.3.4. Evaluation des performances des modèles de régression.....	74
III.4. Techniques de validation des modèles	74
III.4.1. Simple validation.....	75
III.4.2. Validation croisée K-Fold	75
III.4.3. Validation croisée Leave-one-out	76
IV. Conclusion.....	76
Références	76
Chapitre 3 : CONCEPTION ET RÉALISATIONS EXPÉRIMENTALES ET METHODOLOGIE.....	80
I. INTRODUCTION.....	81
II. DISPOSITIF EXPERIMENTAL	81
II.1. Types de conceptions des systèmes nez électroniques	82
II.1.1. Dispositif ouvert	82
II.1.2. Dispositif statique	82
II.1.3. Dispositif dynamique.....	84
II.2. Conceptions et réalisations	85
II.2.1. Choix des capteurs.....	85
II.2.2. Matrice de capteurs.....	88
II.2.3. Alimentation de la matrice des capteurs.....	89
II.2.4. Système d'acquisition de données.....	91
II.2.4.1. Carte d'acquisition de données ADLINK USB-1901	91
II.2.4.2. Carte d'adaptation et de protection.....	92
II.2.4.3. Description du logiciel utilisé LabVIEW	94
II.3. Présentation de l'ensemble du dispositif expérimental	95
II.3.1. Schéma global.....	95
II.3.2. Conditions expérimentales.....	96
II.3.3. Procédure de mesure.....	96
II.3.4. Logiciel de traitement de données	97
III. Conclusion.....	98
Références	98
Chapitre 4 : RESULTATS ET DISCUSSIONS : ETUDE DE CAS DE LA MENTHE COMMESTIBLE	100
Introduction	101
I. Première partie : DETECTION DES INSECTICIDES DANS LA MENTHE COMMESTIBLE	102
I.1. Préparation des échantillons	102
I.2. Réponses temporelles des capteurs	103

I.3. Discrimination de la menthe traitée de celle non traitée.....	105
I.3.1. L'analyse en composantes principales (ACP) à deux dimensions.....	106
I.3.2 Les machines à vecteurs de support (SVM).....	108
I.4. Conclusion	111
II. Deuxième Partie : ETUDE DE CAS DE TRAITEMENT DE LA MENTHE PAR LE MALATHION.....	112
II. 1. La préparation des échantillons	112
II.2. Réponses temporelles des capteurs.....	113
II.3. Discrimination de la menthe traitée de celle non traitée.....	115
II.3.1. Projection de différents types de menthe par ACP à trois dimensions.....	115
II.3.2. Identification de la menthe non traitée à partir de la menthe traitée par la méthode KNN	117
II.4. Prédiction du jour de traitement de la menthe avec le réseau NARX	119
II.5. Comparaison des résultats du réseau NARX avec PLSR et ϵ -SVM	122
II.4. Conclusion	125
Références	126
Chapitre 5 : RESULTATS ET DISCUSSIONS : ETUDE DE CAS DES POMMES DE TERRE	129
Introduction	130
I. Première Partie : IDENTIFICATION DES POMMES DE TERRE TRAITÉES AVEC LES ENGRAIS NPK DE CELLES TRAITÉES TRADITIONNELLEMENT.....	130
I.1. Préparation des échantillons	131
I.2. Réponses temporelles des capteurs	132
I.3. Distinction des pommes de terre selon leurs champs de culture	134
I.3.1. Résultats de l'ACP	134
I.3.2. Résultats du SVM	135
I.4. Conclusion	137
II. Deuxième Partie : DISCRIMINATION DES POMMES DE TERRE A L'AIDE D'UN SYSTEME MULTISENSORIEL ET D'UNE METHODE MICROBIOLOGIQUE	137
II.1. Préparation des échantillons	138
II.2. Analyse microbiologique.....	138
II.2.1. Principe de la méthode de dénombrement de la population microbiologique totale.....	138
II.2.2. Résultats et discussion	139
II.3. Analyse avec le nez électronique.....	140
II.3.1. Préparation des échantillons	140
II.3.2. Résultat de l'analyse multisensorielle	141
II.3.2.1 Réponses du réseau de capteurs et prétraitement	141
II.3.2.2 Résultat de la classification	144

II.4. Conclusion	149
Références	149
CONCLUSION GENERALE	153

INTRODUCTION GENERALE

L'homme est doté de sens qui l'aident à identifier la nature des choses qui l'entourent, et parmi eux l'odorat, qui l'aide à distinguer entre les bonnes et les mauvaises odeurs. Récemment, une grande attention a été accordée à essayer d'imiter ce sens naturel et de l'utiliser au profit de l'humanité dans de nombreux domaines, en particulier celui de l'industrie alimentaire. Auparavant et jusqu'à présent, l'homme prenait ses décisions en fonction de son odorat puis de celui des animaux, en particulier des chiens, mais ces méthodes nécessitent de l'exercice et sont sujettes à la fatigue et à la défaillance en termes de fiabilité. C'est pourquoi l'homme, après ses progrès scientifiques, a eu recours aux analyses chimiques. Ces dernières ne sont pas non plus sans inconvénients, car elles nécessitent des équipements coûteux, des techniciens spécialisés et formés, ainsi qu'un temps relativement important. Par conséquent, l'homme a pensé à simplifier ces tâches de contrôle.

Toujours dans la recherche de la simplification des tâches de contrôle des aliments, les dernières recherches ont abouti à l'imitation du nez humain avec un outil électronique s'appelant le nez électronique ou encore système multi-capteurs. Cet outil est composé d'un réseau de capteurs couplé à des méthodes de reconnaissance de formes. Depuis l'apparition de leur principe de fonctionnement, ces outils ont attiré une grande attention des chercheurs qui tentent de construire des analogues fonctionnels du sens olfactif des mammifères en se basant sur les différents capteurs de gaz.

L'utilisation des nez électroniques, ou ce que nous pouvons l'appeler une « olfaction artificielle », est un domaine prometteur qui a suscité un grand intérêt en recherche et développement. Cependant, il n'y a pas un dispositif convenu qui soit applicable dans tous les domaines, car chaque nez électronique est déterminé selon les capteurs qui le composent en fonction du besoin et de la nature de l'utilisation.

Cette thèse a pour objectif d'étudier et d'évaluer l'efficacité d'un système multi-capteurs (nez électronique) réalisé en laboratoire à l'aide de capteurs commerciaux de gaz à oxyde métallique dans le contrôle qualité des produits agricoles. Ainsi, l'objectif principal de la thèse est d'étudier si le nez électronique est capable de distinguer les échantillons en fonction de

l'odeur qu'ils émettent, ce qui aidera à classer et à identifier la nature des échantillons pour s'assurer de la qualité et de la sécurité des aliments. Après la première étape d'apprentissage, où le nez électronique interagit avec l'odeur et convertit les informations en signaux électriques qui seront analysés par la suite pour former des empreintes olfactives de référence. Le nez électronique peut analyser les composés organiques volatils des différents genres d'échantillons, que leur état chimique soit liquide, gazeux ou solide. Ainsi il peut distinguer les différences et prouver la conformité ou la non-conformité des échantillons.

Ce rapport de thèse se compose de cinq chapitres, le résumé suivant décrit la portée de chacun :

Un premier chapitre sera consacré à la présentation de l'état de l'art sur les odeurs, le système olfactif humain, le nez électronique et quelques méthodes d'analyse des odeurs. Après une étude de l'odeur, le nez électronique sera décrit dans une grande partie en mettant l'accent sur sa composante principale les capteurs de gaz et plus précisément les capteurs de gaz à oxyde métalliques. En dernier, les méthodes physico-chimiques d'analyse des odeurs les plus utilisées seront décrites.

Le deuxième chapitre portera sur l'analyse de donnée et l'apprentissage automatique (en anglais : Machine Learning (ML)) où nous décrirons dans un premier temps les méthodes du prétraitement des données et les différentes techniques de choix des caractéristiques, puis les principaux algorithmes utilisés que nous avons exploité dans nos travaux en l'occurrence, pour la classification des données nous avons : l'Analyse en Composantes Principales (ACP), les Séparateurs à Vaste Marge (SVM, en anglais : Support Vector Machines), le K plus Proches Voisins (KNN, en anglais : K Nearest Neighbors), et pour la régression nous avons : la régression par moindres carrés partielles (PLS, en anglais : Partial least Square), la régression des Séparateurs à Vaste Marge (SVMR, en anglais : Support Vector Machines Regression), et le Réseau Autorégressif Non linéaire à entrées eXogènes (NARX, en anglais : Nonlinear Autoregressive Network with eXogenous inputs).

Le troisième chapitre sera réservé à la conception matérielle qui présentera l'implémentation et la réalisation au sein du laboratoire d'un dispositif multi-capteurs à base de capteurs de gaz à oxyde métallique de commerce. Au préalable, nous présenterons les conceptions possibles pour un nez électronique, leurs avantages et inconvénients pour choisir

la conception la plus adaptée à notre étude. Les composants de notre système seront également décrits en détail.

Le quatrième chapitre sera dédié à une étude de cas réalisé sur la menthe, il révélera tous les résultats expérimentaux obtenus en évaluant dans quelle mesure la menthe est exempte des insecticides en analysant ses composés volatils. Dans une première partie, nous dévoilerons que le système proposé est capable d'identifier la menthe comestible sans insecticides de celles traitées avec deux types des dits produits. La seconde partie, sera consacrée spécialement à un insecticide largement utilisé au monde par les agriculteurs où nous allons prouver encore la faisabilité de l'utilisation de notre outil dans la discrimination de la menthe non traitée de celle traitée avec la prédiction du jour de traitement.

Le cinquième chapitre sera quant à lui consacré à l'étude du cas des pommes de terre. Dans une première partie, nous allons mener une étude pour identifier la source des échantillons des pommes de terre récoltées à partir de deux champs différents, un traité traditionnellement avec le fumier de moutons et d'ânes et un traité avec les engrais NPK (N : Azote, P : phosphore, K : potasse), en s'appuyant sur les résultats des espaces de têtes des échantillons. Dans une deuxième partie, l'étude sera élargie en ciblant trois types de champs : traités traditionnellement avec le fumier de moutons et d'ânes, traités avec les engrais NPK et traités avec le fumier de poulet, en utilisant aussi une analyse microbiologique pour la comparaison et la confirmation de l'utilité de l'outil proposé.

Tous les tests numériques ont été réalisés sur Matlab 2019a sous licence ENSA.

A la fin nous achèverons ce mémoire avec une conclusion dans laquelle nous présenterons un résumé des discussions sur les résultats globaux atteints dans le cadre de ce travail avec des perspectives sur les possibles applications du nez électronique dans de futures recherches en agroalimentaire.

Chapitre 1 : METHODES D'ANALYSE DES ODEURS ET CAPTEURS DE GAZ

I. INTRODUCTION

Les signaux chimiques jouent un rôle très important dans tous les aspects de la vie de presque tous les êtres vivants, y compris les humains. Ces signaux contiennent des informations importantes sur de nombreux domaines tels que la nutrition, la reproduction, la reconnaissance de parentèle et la détection de conditions potentiellement dangereuses telles que les prédateurs, les volcans, les feux ou les aliments toxiques, etc.

L'odorat humain est un sens analytique naturel de très grande importance, simple dans sa fonction habituelle mais très complexe à reproduire. Pour surmonter certains problèmes et pour une étude plus approfondie des odeurs émises par différents matériaux, qui contiennent une grande source d'informations précieuses sur leurs structures et leurs composants, les chercheurs utilisent des outils physico-chimiques tels que la chromatographie en phase gazeuse et la technologie de spectrométrie de masse. Ces outils, même s'ils sont sensibles et donnent de bons résultats, sont chers, encombrants, difficiles d'utilisation et nécessitent des experts qualifiés pour leurs manipulations. Ainsi, dans le cadre de la recherche et du développement continu, les chercheurs scientifiques sont intéressés d'inventer des appareils électroniques tentant d'imiter le travail du nez humain connu sous le nom de nez électronique. Cet outil est composé principalement d'un ensemble de capteurs de gaz ainsi que d'un mécanisme de reconnaissance de formes.

Dans ce premier chapitre, nous présenterons l'odeur, ses types et caractéristiques, puis nous verrons le nez biologique humain, son architecture en liaison avec son principe de fonctionnement et les facteurs influençant la perception des odeurs. Ensuite, nous présenterons le nez artificiel connu sous le nom de nez électronique, une dissertation détaillée sera donnée sur son historique et son composant principal qui est le capteur de gaz en mettant l'accent sur les capteurs de gaz à oxydes métalliques. Vers la fin de ce chapitre, nous présenterons quelques méthodes physico-chimiques les plus exploitées pour l'analyse des odeurs qui peuvent être utilisées seules ou couplées à savoir la chromatographie en phase gazeuse (CG), la spectrométrie de masse (SM), la chromatographie en phase gazeuse couplée à la spectrométrie de masse (CG-SM) ou encore la chromatographie en phase gazeuse couplée à l'Olfactométrie (CG-O).

II. LES ODEURS

II.1. Définition

L'odeur désigne généralement la sensation perçue provoquée par des mélanges simples ou complexes de composés volatils présents dans l'air à des concentrations différentes stimulant l'organe olfactif. Elle a été définie par Brancher et coll. [1] comme « une sensation créée par l'interaction d'espèces chimiques volatiles inhalées par le nez, telles que les composés soufrés (par exemple les sulfures, les mercaptans), les composés azotés (par exemple l'ammoniac, les amines) et les composés organiques volatils (par exemple les esters, les acides, aldéhydes, cétones, alcools) ». Après avoir cru que l'être humain peut distinguer 10000 odeurs dans la littérature profane et scientifique, il a récemment été estimé qu'il peut discriminer au moins 1 billion de stimuli olfactifs sur la base des résultats des tests psychophysiques menés par Bushdid et coll. [2]. Cette merveilleuse conception du nez humain est la base des études de l'analyse des odeurs.

Les odeurs reconnaissables sont le résultat d'un mélange compliqué de plusieurs molécules odorantes qui sont généralement des minuscules molécules hydrophobes, chacune ayant une concentration différente. La majorité des parfums et arômes naturels sont des combinaisons complexes d'espèces chimiques, avec des dizaines, voire des centaines d'ingrédients. Comme exemple, la saveur du café qui est extrêmement complexe, contient plus de 1000 composés qui affectent les perceptions sensorielles de la muqueuse buccale et nasale [3].

II.2. Rôle de l'odeur

L'odorat des êtres vivants joue un rôle majeur dans leur vie quotidienne. Les animaux utilisent ce sens pour s'accoupler, pour identifier les parents, pour préserver leur vie en distinguant l'odeur des prédateurs ou encore pour chasser en traquant l'odeur des proies. De même pour l'être humain dans le passé, l'odorat l'a toujours aidé à distinguer les aliments, à découvrir les incendies et à se tenir à l'écart des dangers tels que les volcans et les prédateurs.

Après le développement de l'homme, c'est-à-dire à notre époque, l'odorat joue toujours un rôle majeur dans l'industrie et le commerce, car il est utilisé pour attirer les consommateurs en ajoutant une odeur aux produits de cuisine, aux vêtements et aux produits de soins de santé personnels, il est aussi utilisé dans la sécurité de l'homme avec le gaz du propane pour découvrir ses fuites en plus d'autres utilisations.

Malgré l'importance de ce sens dans la vie humaine, il est moins puissant que son homologue chez certains animaux, ce qui explique l'utilisation de chiens de chasse ou d'oiseaux canaris comme système pour avertir les mineurs de la présence de gaz explosifs.

II.3. Propriétés des odeurs

Les qualités communes des odeurs comprennent :

- En présence d'une odeur forte, une odeur faible n'est pas détectée.
- La force constante de l'odeur fait qu'une personne perd conscience de celle-ci, qui n'est remarquée que lorsqu'elle varie en intensité.
- Les odeurs comparables peuvent être trouvées dans des substances ayant des compositions chimiques similaires ou distinctes. La nature et la force de l'odeur peuvent s'altérer lorsqu'elle est diluée.
- Une personne peut sentir une odeur de loin.
- Un mélange des odeurs de force similaire pour créer un mélange dans lequel un ou les deux composants peuvent être méconnus.
- La fatigue due à une odeur peut ne pas altérer la perception d'autres odeurs différentes, mais elle interfère avec la perception d'odeurs comparables.
- Les goûts et les dégoûts sont souvent basés sur le lien de l'odeur avec des souvenirs agréables ou mauvais.
- Une odeur inconnue est plus susceptible qu'une odeur familière de déclencher une reconnaissance d'odeur.
- L'odeur est transportée sous le vent.
- De nombreux animaux ont un odorat plus puissant que les humains.

II.4. Paramètres des odeurs

Les paramètres des odeurs [1] listés ci-dessous peuvent être définis :

- Seuil olfactif : la plus petite concentration d'un produit chimique odorant requise pour induire une réponse spécifique chez l'homme est connue sous le nom de seuil olfactif (concentration). Ils sont fréquemment donnés en pourcentage, par exemple 50 % ou 100 %, selon le nombre de personnes qui détectent l'odeur. Étant donné que l'odeur des mélanges odorants ne peut pas être caractérisée en ajoutant simplement les effets

d'odorants uniques, il est essentiel de souligner que les seuils d'odeur ne sont pas additifs.

- **Intensité** : c'est une mesure décrivant l'intensité d'un stimulus olfactif, et elle peut être liée à la concentration de l'odeur par des formules mathématiques spécifiques. Dans notre vie quotidienne, elle peut être exprimé par des expression allant de non perceptible à extrêmement forte.
- **Diffusivité** : une odeur n'est ressentie que lorsqu'une molécule odorante atteint l'épithélium olfactif ; par conséquent, la volatilité (représentée par la pression de vapeur) est un paramètre clé pour déterminer la capacité d'une substance chimique à produire une odeur. L'indice d'odeur est défini comme le rapport de la pression de vapeur d'un composé odorant à son seuil olfactif.
- **Qualité** : elle peut être déterminée en utilisant des classes d'odeurs ou des échelles nominales de gamme de mesure (catégorie), comme la roue des odeurs proposée par Fisher et coll. en 2018 [4].
- **Tonalité hédonique** : la tonalité hédonique est une échelle qui peut être utilisée pour exprimer le caractère agréable ou désagréable d'une odeur, allant de -4 (très désagréable) à + 4 (extrêmement agréable), avec 0 représentant une odeur qui n'est ni agréable ni désagréable.

III. LE NEZ HUMAIN

III.1. Architecture du système olfactif humain

Les qualités d'une odeur sont liées à la structure, la taille et les caractéristiques polaires des molécules qui les composent. Les matériaux doivent être en suspension dans l'air et finement séparés pour déclencher l'odorat. Les liquides, par exemple, ne peuvent pas être sentis jusqu'à ce qu'ils se vaporisent. La volatilité semble être une condition requise mais pas suffisante pour stimuler l'odorat. Étant donné que les composants de l'odeur doivent être collectés par la membrane muqueuse des narines avant de pouvoir être stimulés, donc la solubilité est nécessaire.

Le système olfactif humain est très complexe et n'est pas encore entièrement compris [5]. L'anatomie du système olfactif des êtres humains dévoile sa composition de trois sous-systèmes principaux : l'épithélium olfactif, le bulbe olfactif et le cortex olfactif (système limbique) [6]. La figure 1-1 donne un aperçu sur le système olfactif chez un être humain en mentionnant les positions des sous-systèmes qu'on va développer par la suite.

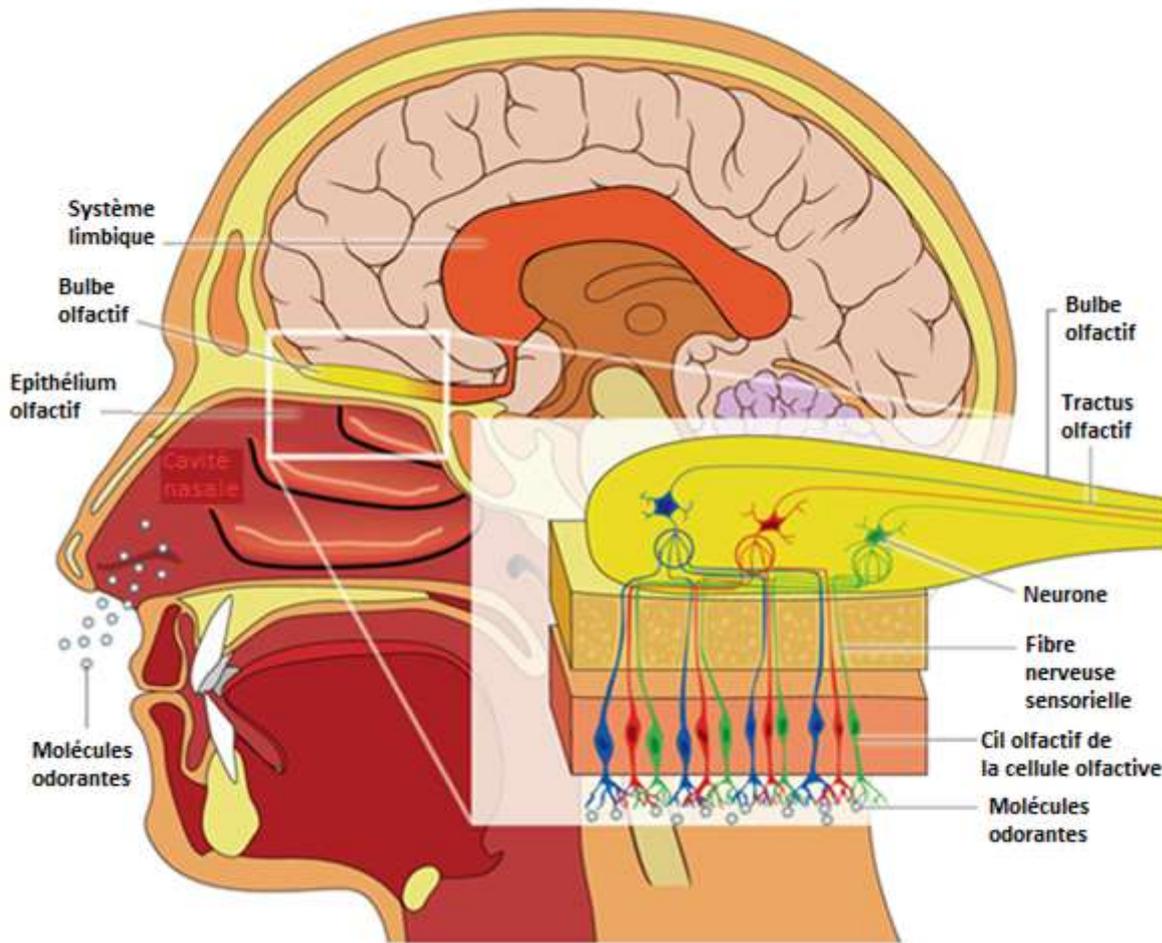


Figure 1-1 Système olfactif humain ; Illustration de Patrick J. Lynch, illustrateur médical

III.1.1. L'épithélium olfactif

Le processus de la reconnaissance de l'odeur commence après un simple reniflement, dans lequel les molécules d'odeurs émises par n'importe quelle substance stimulent le système olfactif et l'échantillon contenant de l'odeur est transféré à l'épithélium dans lequel se trouvent les récepteurs des odeurs qui se stimulent chimiquement. L'épithélium olfactif reçoit une odeur à travers un réseau de voies nasales, appelées narines. Les cellules réceptrices, les cellules de soutien et les cellules basales constituent la majorité de l'épithélium olfactif. Les cellules réceptrices recouvrent une seule couche de cellules basales et sont placées en mosaïque entre les cellules de soutien. Les odorants traversent une couche de mucus pour se connecter aux membranes chimiquement sensibles des cellules réceptrices olfactives de l'épithélium. La première étape de la création du signal est supposée avoir lieu dans un récepteur olfactif, et cette réaction semble dépendre des récepteurs des neurotransmetteurs de la protéine G. Le nombre de divers types de protéines réceptrices est limité (environ 1 000), mais le nombre de

cellules réceptrices est vaste (environ 100 millions) [7]. En conséquence, les cellules réceptrices individuelles ont des sensibilités qui se chevauchent, c'est-à-dire que les récepteurs ne sont pas intrinsèquement spécifiques à des substances odorantes uniques et n'ont pas de sélectivité imposée par le mucus. En conséquence, les étapes ultérieures du système olfactif doivent reconnaître le modèle d'activité produit par cet ensemble de cellules réceptrices qui portent des protéines réceptrices spécifiques aux molécules d'une odeur donnée.

III.1.2. Bulbe olfactif

Le bulbe olfactif (la station de traitement du signal), composé de trois couches : glomérules, cellules mitrales et couche granuleuse [8], il est le premier relais des apports olfactifs. Les cellules mitrales traitent les signaux électriques et les envoient au cortex olfactif via la couche granuleuse, qui crée une couche épaisse, vers le cortex olfactif. Le cerveau envoie des messages de rétroaction aux cellules mitrales sur l'état émotionnel actuel. Un processus superficiel des cellules granulaires se termine dans la couche plexiforme extérieure, établissant de nombreuses connexions avec les dendrites latérales des cellules mitrales. Les cellules granulaires n'ont pas de véritable axone et leurs processus, qui sont parsemés d'innombrables épines microscopiques, ressemblent à des dendrites dans leurs minuscules caractéristiques structurelles. Dans le système olfactif, la couche granuleuse offre une spécificité. En conséquence, les signaux olfactifs sont soumis à un traitement poussé à ce niveau, ce qui minimise le bruit du signal et l'amplifie, augmentant ainsi à la fois la sensibilité et la sélectivité.

III.1.3. Cortex olfactif

Le cortex olfactif du cerveau est le niveau ultime du système olfactif des mammifères. Bien que le cortex olfactif primaire soit encore une partie du cerveau encore peu étudiée, il est connu pour être majoritairement constitué de connexions associatives entre les cellules pyramidales et le tractus olfactif latéral [9]. La sortie du bulbe olfactif est dirigée vers le cortex olfactif, qui est situé sur les surfaces basale et latérale du cerveau antérieur. Il sert de contenu - mémoire adressable pour l'association de stimuli olfactifs avec des traces mnésiques de stimuli olfactifs précédents. En conséquence, les données olfactives sont transformées en une expérience sensorielle cohérente, ainsi que la création et l'évocation de souvenirs de noms, de lieux et de sentiments, entre autres, qui relie l'expérience actuelle à des expériences antérieures avec le mélange. Le cerveau humain comprend environ 10^{11} de neurones et 10^4 interconnexions par neurone, ce qui en fait un ordinateur extrêmement complexe, non linéaire

et parallèle [10]. Il a la capacité d'organiser les neurones de manière à pouvoir exécuter des calculs tels que l'odorat. L'identification plusieurs fois plus rapide que l'ordinateur numérique le plus rapide actuel. De nombreuses tentatives ont été entreprises au cours des dix dernières années pour utiliser des modèles de réseaux de neurones artificiels pour représenter le cortex olfactif dans le cerveau.

III.2. Facteurs influençant la perception des odeurs

Les réponses individuelles à l'exposition aux odeurs sont aussi très personnelles et subjectives. Les facteurs suivants sont les plus ennuyant à la perception olfactive humaine à l'évaluation d'une odeur :

- Les antécédents personnels d'exposition à des substances odorantes identiques ou comparables.
- Le lien personnel avec l'établissement responsable de la source de l'odeur.
- Les souvenirs olfactifs.
- Les impressions psychologiques d'événements liés à ces épisodes d'exposition.
- L'humeur.
- La technique d'adaptation personnelle.
- Les caractéristiques physiques telles que l'âge, la santé, le sexe et la génétique.

III.3. Limite d'utilisation du nez humain

Le système olfactif humain est soumis à des facteurs limitant son utilisation car il ne peut pas être utilisé comme un dispositif de détection dans de nombreuses situations qui mettent la vie humaine en danger surtout dans l'industrie qui contient des composants nocifs, en addition du problème de la cessation de la perception de la même odeur après sa première reconnaissance.

En plus il a été démontré que le nez humain est un incapable de reconnaître les composants d'un mélange qui dépasse 4 odorants [11] sans oublier la variation de seuil de perception consciente des odeurs.

IV. LE NEZ ELECTRONIQUE :

IV.1. Définition

Les nez électroniques ou les systèmes multi-capteurs de gaz sont des appareils électroniques qui essaient d'imiter le principe de fonctionnement du système olfactif humain pour la reconnaissance des odeurs [12]. Ils peuvent être utilisés pour augmenter ou remplacer les groupes d'experts humains dans certaines circonstances. Les premiers travaux sur la mise au point d'un appareil spécialement destiné à la mesure des odeurs remontent aux années cinquante du siècle dernier, mais ces appareils n'ont attiré l'attention des chercheurs que dans les années quatre-vingt où le vrai travail a commencé pour les développer. Les nez électroniques se composent principalement d'un ensemble de capteurs chimiques qui agissent comme un récepteur réagissant aux odeurs qui les pénètrent et d'un système de reconnaissance de formes.

IV.2. L'historique du nez électronique :

Comme mentionné précédemment, les travaux sur les premiers nez électroniques remontent au début des années cinquante du siècle dernier, plus précisément en 1954 par Hartman qui a présenté le premier prototype d'un appareil électrique capable de détecter les odorants [13]. Comme éléments de détection, le système utilisait des microélectrodes polarisées. En 1961, Moncrieff [14] a examiné plusieurs matériaux de revêtement qui interagissaient différemment avec les odorants en utilisant une thermistance comme dispositif de transduction. Il a déclaré qu'en combinant une variété de capteurs avec différents revêtements, le spectre de détection pourrait être élargi, et donc un grand nombre d'odeurs pourrait être distingué. Seiyama et coll. [15] ont créé un capteur de gaz à couche mince semi-conductrice en 1962. Les changements de conductivité électrique ont été utilisés comme base pour leur méthode de détection de gaz. En 1964, Hartman revient cette fois-ci avec Wilkens [16] pour produire le premier appareil électronique capable de distinguer les odeurs en basant leurs recherches sur le phénomène des réactions d'oxydo-réduction de l'odeur à l'électrode. En 1965, Buck et coll. [13] ont publié des travaux similaires à ce qu'il a fait Seiyama et ses collègues en 1965. Dravnieks et Trotter [17] ont créé un détecteur de vapeur basé sur la modulation thermique du potentiel de contact la même année. En 1967, Shaver [18] a présenté une méthode pour augmenter la sensibilité d'un détecteur de gaz à l'oxyde de tungstène en ajoutant un matériau catalytique comme le platine. Taguchi a développé l'année suivante les premiers capteurs de gaz à semi-conducteur à oxyde métallique (MOS) à usage résidentiel et industriel,

utilisant l'oxyde d'étain comme matériau de revêtement sensible, qu'il a breveté en 1971 [19]. Figaro Engineering Inc., sa société, est devenu le premier fabricant de capteurs de gaz MOS.

Cependant, une décennie plus tard, le concept d'un nez électronique intelligent capable de classer les odeurs refait surface dans les travaux de Persaud et Dodd en 1982 à l'Université de Warwick au Royaume-Uni [20] et dans ceux de Ikegami et ses partenaires entre 1985 et 1988 chez Hitachi au Japon [21].

L'expression « nez électronique », qui a été créé par Gardner et Bartlett en 1988, est également réapparu en vigueur. Cette fois dans des événements et conférences comme la session de l'atelier avancé de NATO (the North Atlantic Treaty Organization) sur le traitement des informations sensorielles chimiques qui a été consacrée au thème de l'olfaction artificielle et à la conception du système olfactif artificiel en 1989. En plus, En 1991, la première conférence consacrée uniquement au thème du nez électronique a eu lieu [22]. Ensuite les événements se sont succédés jusqu'à ce jour.

À présent, cet appareil suscite toujours un grand intérêt en recherche et développement, alors qu'il est commercialisé par plusieurs sociétés.

IV.3. Principe de fonctionnement du nez électronique

Inspiré du nez humain, le principe de fonctionnement du nez électronique tente d'imiter celui du nez humain. Premièrement, le mélange de gaz entre en contact avec le réseau de capteurs qui joue le rôle du récepteur, ce contact entraîne une variation dans une grandeur électrique comme la résistance, la tension, le courant ou la fréquence, en fonction des catégories des composants du réseau de capteurs. Les signaux bruts engendrés par les capteurs sont enregistrés pour un prétraitement afin de comprendre correctement ces changements physiques, puis traités pour les numériser afin de former une empreinte digitale (signature olfactive) qui sert comme référence pour classer et distinguer les odeurs. En se reposant sur la comparaison des empreintes digitales, cette méthode d'identification de l'odeur est loin des méthodes purement analytiques.

La figure 1-2 [23] représente les composants du système olfactif humain et leurs homologues dans le nez électronique.

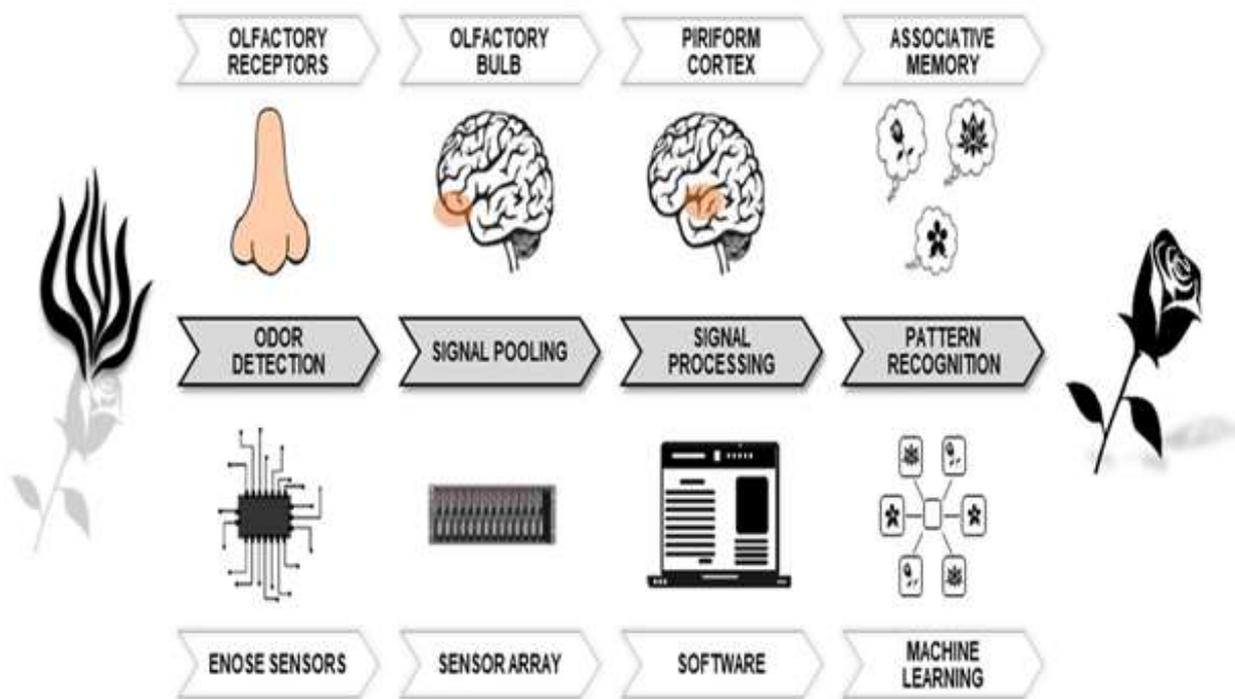


Figure 1-2 Comparaison schématique de la technologie du nez électronique et du système olfactif [23]

IV.4. Capteurs de gaz

IV.4.1. Les différents types de capteurs de gaz

Les capteurs de gaz sont utilisés pour détecter et identifier les composés volatils, que ce soit dans les matériaux ou dans l'air ambiant, en fonction de la concentration des gaz qui les composent. Pour cette raison, des recherches et des études ont été menées en permanence pour développer différents types de capteurs, ces derniers sont classés en fonction des différents phénomènes physiques sur lesquels ils se reposent, parmi ces capteurs, on trouve : les capteurs résistifs à oxydes métalliques à semi-conducteurs (MOX), les capteurs capacitifs, les capteurs électrochimiques, les capteurs catalytiques, les capteurs optiques, les capteurs acoustiques, les capteurs infrarouges ...etc.

Dans le tableau 1-1, un résumé des différents capteurs, leurs matériaux sensibles et leurs principes de fonctionnement.

Tableau 1-1 Différents types des capteurs, leurs matériaux sensibles et leurs principes de fonctionnement

Type de capteur	Matériau sensible	Principe de détection
Oxydes métalliques à semi-conducteurs (MOX)	Oxydes métalliques semi-conducteurs dopés (SnO ₂ , GaO)	Changement de résistance
Polymères conducteurs	Polymères conducteurs modifiés	Changement de résistance
Capteurs acoustiques : QMB, SAW, BAW	Couches de film organiques ou inorganiques	Changement de masse (décalage de fréquence)
Capteurs électrochimiques	Électrolytes solides ou liquides	Changement de courant ou de tension
Capteurs optiques	Photodiode sensible à la lumière	Changements optiques
Capteurs infrarouges	Détecteur sensible aux infrarouges	Absorption des rayonnements infrarouges

Utilisé dans ce travail, notre étude bibliographique par la suite portera davantage sur les capteurs de gaz à oxyde métallique (MOX).

IV.4.2. Capteurs résistifs à oxydes métalliques semi-conducteurs (MOX)

Les capteurs résistifs à oxydes métalliques semi-conducteurs (MOX) sont le type le plus répandu et utilisé grâce à leurs multiples avantages comme le faible coût de production et de commercialisation, la fiabilité, la miniaturisation, la faible limite de détection (< niveaux de ppm), l'usage facile et intégration dans l'électronique moderne. Bien qu'ils présentent certains inconvénients comme la consommation élevée d'énergie, la sensibilité à l'humidité et la faible sélectivité, ils suscitent toujours un grand intérêt des chercheurs.

IV.4.2.1. Historique

Les premiers capteurs de gaz semi-conducteur chimiorésistant ont été commencés, au début des années soixante, par le chercheur Seiyama [15] en se basant sur des études antérieures de Brattain et Bardeen [24] qui ont confirmé l'influence de l'interaction des composés volatils avec les surfaces des semi-conducteurs sur la conductance électrique des semi-conducteurs. Seiyama avait aperçu que des couches minces de ZnO chauffées à $\sim 300^{\circ}\text{C}$ dans l'air, a une conductivité électrique très sensible à la présence de traces de gaz réactifs dans l'air. Tandis que Taguchi, a trouvé des propriétés similaires pour le SnO₂ fritté, avec l'avantage supplémentaire d'une plus grande stabilité en 1962 [25]. Mais le problème avec ces premiers capteurs résidait dans les transformations chimiques irréversibles qui ont connu les semi-conducteurs utilisés ce qui ne permet pas la répétabilité. Après des recherches de Heiland et Lamatsch [26], Bielanski et coll. [27] et Seiyama et coll. [28] il s'est avéré que les oxydes métalliques sont les plus appropriés grâce à la liaison réversible de l'oxygène. Un pas important a été réalisé lorsque Taguchi a breveté sa fabrication d'un capteur semi-conducteur à base d'oxyde métallique en dioxyde d'étain (SnO₂) en 1971 et dès cette année, il a commencé à les commercialiser jusqu'à aujourd'hui [19]. En plus de la fameuse société Figaro de Taguchi, de nombreuses entreprises proposent ce type de capteurs, comme FIS, MQ, e2v, MicroSense, Applied-Sensors ...etc.

Un grand nombre de matériaux d'oxydes métalliques sont utilisés dans les couches actives des capteurs de gaz, cependant, seuls les dioxydes d'étain SnO₂, les trioxydes de tungstène WO₃, les oxydes de zinc ZnO et les dioxydes de titane TiO₂ sont commercialement utilisés.

IV.4.2.2. Structure d'un capteur à oxydes métalliques

En général, l'élément sensible du capteur de gaz MOX est composé de trois éléments principaux déposés les uns sur les autres, en plus de la couche sensible, un élément de chauffage est aussi présent du fait que les capteurs à oxydes métalliques fonctionnent à température élevée (typiquement entre 300°C et 500°C), en plus d'un substrat sur lequel on trouve les électrodes de contact. La figure 1-3 (a) présente la structure la plus ancienne, proposée par Taguchi (capteur Figaro série TGS8xx).

La couche sensible est déposée sur un tube en céramique poreux de diamètre de 1 mm et de longueur de 3 mm. L'élément de chauffage est un fil métallique bobiné à l'intérieur du tube. Les contacts électriques de la couche sensible sont fixés aux deux extrémités du tube avec

une colle conductrice. Les progrès de la microélectronique permettant la miniaturisation des composants électroniques et par conséquent une nouvelle génération de capteurs de gaz est apparue. La figure 1-3 (b) présente la structure d'un μ -capteur de gaz, réalisée en technologie silicium. Ces microstructures permettent de réduire la consommation électrique. La miniaturisation permet d'associer plusieurs capteurs sur un même substrat et ainsi de réaliser

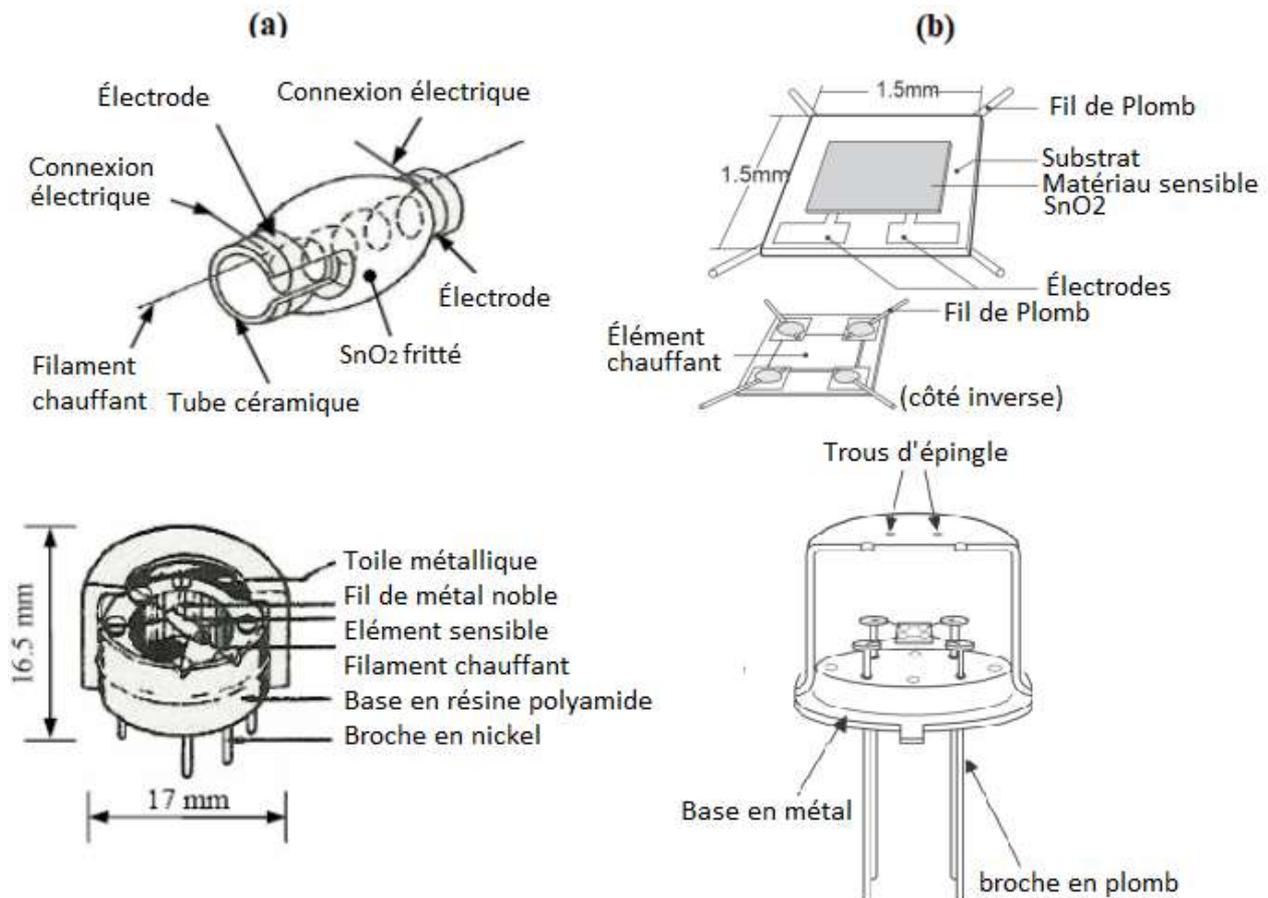


Figure 1-3 Architecture d'un capteur de gaz à oxyde métallique de la société Figaro, (a) la structure la plus ancienne [29] et (b) la structure récente [30]

un réseau de capteurs.

IV.4.2.3. Principe de fonctionnement

Le principe de fonctionnement de ce genre de capteurs repose sur le changement de la conductance de la couche sensible semiconductrice en présence d'un gaz. En effet, en présence d'un gaz, un phénomène d'adsorption survient dans l'interaction des molécules de gaz avec la couche sensible et ce qui conduit à certains changements physiques, les plus courantes en termes de mesurandes électriques (courant, résistance et tension).

Dans le cas des oxydes métalliques, il est possible de distinguer deux phénomènes intervenant dans l'adsorption des gaz : la physisorption et la chimisorption.

La physisorption (l'adsorption physique) fait intervenir des forces d'origine électrostatique de type Van der Waals avec une énergie de liaison inférieure à 0.1 eV, tandis que la chimisorption fait intervenir des forces de liaison pour un ordre de grandeur de 1 eV, impliquant une combinaison chimique de la molécule et du solide.

Le matériau de détection peut être un semiconducteur de type-n ou de type-p et le mécanisme de réaction dans ces deux matériaux sont opposés. Dans le cas de semiconducteur de type-n, le matériau est chauffé par le réchauffeur à plusieurs centaines de °C. Le flux d'électricité à l'intérieur du matériau de détection est déterminé par le nombre d'électrons libres. Lorsque le matériau de détection est dans l'air propre, l'oxygène de l'air est adsorbé sur le matériau de détection et attire les électrons libres, provoquant ainsi une augmentation de la résistance du capteur, ce qui conduit à une diminution de flux d'électrons à l'intérieur du matériau de détection.

En présence de gaz réducteurs, le gaz réducteur réagit avec l'oxygène adsorbé et les électrons retenus par l'oxygène adsorbé sont libérés dans le matériau de détection. En conséquence, la résistance du capteur diminue, permettant plus de courant. Au fur et à mesure que la concentration du gaz réducteur augmente, la résistance du capteur diminue davantage, permettant un courant plus important.

En présence de gaz oxydants, leur atome d'oxygène extrait les électrons de la couche plutôt que de les libérer conduisant ainsi à une résistance accrue. En effet ce qui se passe c'est que la zone de déplétion s'épaissit du fait de la diminution de la concentration du porteur ce qui mène à l'augmentation de la résistance.

Tableau 1-2 Mécanisme de réaction dans les capteurs de gaz à oxyde métallique

Type de matériau	Gaz oxydants	Gaz réducteurs
Type-n	Augmentation de la résistance	Diminution de la résistance
Type-p	Diminution de la résistance	Augmentation de la résistance

Il convient de noter qu'il existe une certaine corrélation entre la résistance du capteur et la réduction de la concentration de gaz dans l'air, cette corrélation est appelée caractéristiques de sensibilité.

IV.4.3. Caractéristiques d'un capteur de gaz à oxyde métallique (MOX)

Les performances d'un capteur de gaz à oxyde métallique peuvent être évaluées selon plusieurs critères qui déterminent ses caractéristiques, les plus essentiels sont les 3S : la sensibilité, la stabilité et la sélectivité.

IV.4.3.1. Sensibilité

La sensibilité est l'une des caractéristiques les plus importantes et elle est décisive dans le choix des capteurs de gaz. En effet, le changement de la résistance de la couche sensible à l'interaction avec les gaz reflète la concentration et la diffusion des gaz. Par conséquent, la sensibilité peut être définie comme le rapport entre la résistance dans l'air et la résistance après interaction avec le gaz cible. Elle est exprimée selon le type de matériau semi-conducteur de la couche sensible par :

$$- S = R_{air} / R_{gaz} \quad (\text{Type-n})$$

$$- S = R_{gaz} / R_{air} \quad (\text{Type-p})$$

La sensibilité peut être exprimée en pourcentage :

$$- S (\%) = \left(\frac{R_{air} - R_{gaz}}{R_{air}} \right) * 100 \quad (\text{Type-n})$$

$$- S (\%) = \left(\frac{R_{gaz} - R_{air}}{R_{gaz}} \right) * 100 \quad (\text{Type-p})$$

IV.4.3.2. Sélectivité

La sélectivité est la capacité d'un capteur à reconnaître sélectivement un seul et exact gaz dans un atmosphère en coexistence avec d'autre gaz. Et comme la sélectivité pour un gaz exact est parfois hautement souhaitable, les capteurs de gaz à oxyde métallique ont l'aptitude à détecter une combinaison des gaz plutôt qu'un seul, et c'est principalement en raison de la structure de l'élément sensible. C'est le handicap majeur des capteurs de gaz à oxyde métallique, de nombreuses méthodes et applications ont été mises en œuvre pour résoudre ce problème. Pour notre cas et dans les études sur la caractérisation des odeurs, ce problème se résout par

l'utilisation de la sensibilité croisée qui s'articule sur l'utilisation d'une matrice de plusieurs capteurs de gaz de types différents pour extraire la signature globale du mélange d'odeurs étudié.

La sélectivité (Q) du capteur de gaz à oxyde métallique peut être exprimée par :

$$Q = \frac{\textit{la sensibilité du capteur aux gaz interférents}}{\textit{la sensibilité du capteur au gaz cible}}$$

IV.4.3.3. Stabilité

La stabilité est la capacité d'un capteur de gaz à reproduire les mêmes réponses dans le temps, cela inclut en plus de la réponse, la conservation de la sensibilité, la sélectivité envers les mêmes gaz et le temps de récupération. Le vieillissement de ces capteurs est engendré principalement par leurs conditions de travail surtout qu'ils nécessitent un chauffage continu et à des degrés très élevés.

En addition aux « 3S », d'autres paramètres peuvent être utilisés pour évaluer les performances d'un capteur de gaz :

- La limite de détection : c'est la valeur minimale de concentration du gaz cible que peut détecter le capteur de gaz, elle est exprimée en ppm (partie par million) ou ppb (partie par milliard).
- Le temps de récupération : c'est le temps de repos nécessaire pour un capteur pour régénérer les mêmes résultats. Autrement dit, c'est le temps nécessaire pour libérer 90% de la couche sensible du capteur après avoir été exposé à un gaz cible.

IV.4.4. Facteurs affectant les capteurs de gaz à oxyde métallique (MOX)

Les capteurs à oxydes métalliques peuvent être influencés par une variété des facteurs qu'on peut classés selon deux catégories, une première liée à la conception interne du capteur comme le matériau avec lequel la couche sensible est conçue (ses propriétés et sa microstructure) et son dopage. La deuxième est liée aux conditions environnementales dont ces capteurs vont être utilisés et c'est ce que nous allons développer par la suite puisque dans ce présent travail nous n'avons pas fabriqué des capteurs tout en se limitant juste à utiliser ceux du commerce prêt à exploiter.

IV.4.4.1. Température

La température de la couche sensible joue un rôle primordial dans le fonctionnement des capteurs à oxyde métallique, car elle affecte directement la sensibilité du capteur pour son rôle dans la stimulation des réactions d'oxydo-réductions entre la couche sensible et le gaz cible. Plusieurs études ont rapporté que la sensibilité des capteurs augmente à mesure qu'ils

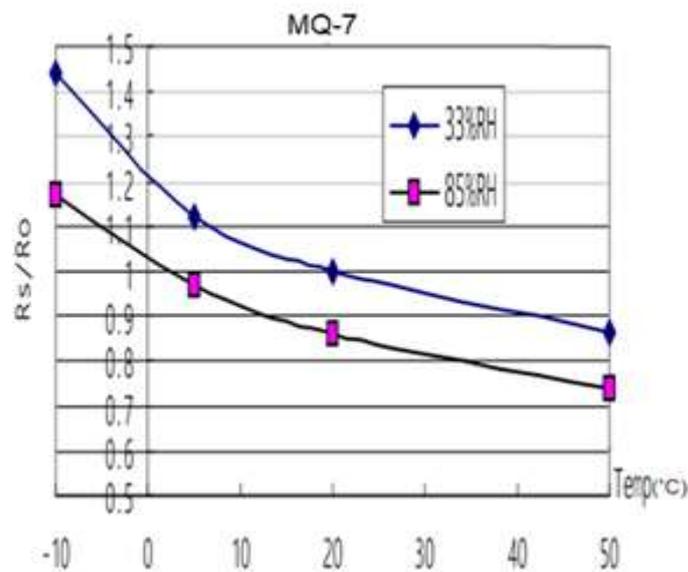


Figure 1-4 Influence de la variation de la température sur la réponse du capteur MQ-7 de commerce à oxyde métallique SnO₂

approchent du degré optimal et diminue lorsqu'on s'en éloigne, et le degré de sensibilité varie en fonction de la température de la couche sensible et en fonction de la nature des différents gaz.

La figure 1-4 montre la dépendance typique du capteur à oxyde métallique de commerce MQ-7 à la température et l'humidité à une concentration du monoxyde de Carbon (CO prédéfinie à 100ppm prise à partir de sa datasheet.

La température du capteur peut être affectée par la température de l'effluent gazeux. Par conséquent, elle doit être contrôlée et stable pendant les mesures. Cela nécessite également d'éviter un débit de gaz excessif et/ou variable.

IV.4.4.2. Humidité

En plus de la température, l'humidité est l'un des facteurs affectant les réactions d'oxydoréduction qui se déroulent à la surface du matériau sensible des capteurs. L'humidité influence sur la sensibilité du capteur car il réagit comme un gaz réducteur ce qui conduit à

impacter les performances du capteur et fausser les mesures. L'influence de l'humidité varie selon deux possibilités :

- La première est l'adsorption des molécules d'eau par la couche sensible ce qui conduit à la réduction du nombre de sites d'adsorption infectant ainsi l'adsorption des espèces oxygénées à la surface du semi-conducteur.
- La deuxième est la présence de l'humidité qui peut se comporter comme un donneur d'électrons ce qui conduit à la modification de la résistance de la couche sensible.

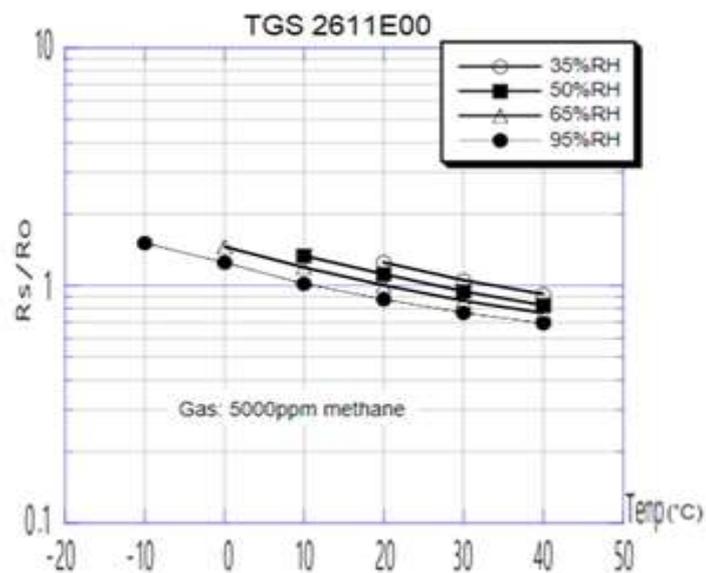


Figure 1-5 Influence de la variation de l'humidité et la température sur la réponse du capteur TGS 2611E00 du commerce

La figure 1-5 représente des réponses du capteur à oxyde métallique de commerce TGS 2611E00 de la société Figaro au gaz CH₄ (méthane) à une concentration stable de 5000ppm avec variation d'humidité de 35% à 95% et de température de -10 à 50 °C prise de sa datasheet.

IV.4.5. Critères de sélection d'un capteur de gaz à oxyde métallique

Lors du choix d'un capteur de gaz à oxyde métallique, il y a quelques critères à garder présents à l'esprit, comme :

- La résolution c-à-d la plus petite valeur qui peut générer un incrément du capteur ;
- Les conditions environnementales toléré pour le bon fonctionnement car on trouve des limites de température/humidité qui sont fréquemment définies ;
- Le coût et la disponibilité sur le marché ;

- La précision ;
- La plage maximale de mesure du capteur ;
- L'étalonnage qui est nécessaire pour la plupart des équipements de mesure car les mesures changent avec le temps ;
- La reproductibilité—La lecture variable est mesurée à plusieurs reprises dans le même réglage ;
- La consommation d'énergie.

IV. 4.6. Champs d'applications

Les systèmes de nez électroniques à base de capteurs de gaz sont conçus et fabriqués pour être utilisés afin d'aider les personnes dans de nombreux domaines et aspects de la vie dans lesquels il leur est difficile de prendre des décisions. Ils ont été utilisés dans de nombreux domaines différents mais l'agroalimentaire reste le plus grand marché pour ces outils. Dans cette industrie, des mesures d'assurance qualité sont nécessaires, ces outils sont utilisés pour inspecter le type et la qualité des ingrédients, superviser le processus de production et superviser toutes les procédures liées à la durée de conservation du produit. De cet effet, ils sont utilisés pour prévenir les fraudes des consommateurs en confirmant les ingrédients et les normes de contenu, et en effectuant des évaluations de contrôle de qualité pour déterminer la cohérence de la marque et du produit [33]–[36].

Dans l'agriculture où ils sont utilisés pour assurer la maturité des cultures et la sécurité alimentaire contre la détérioration, la pourriture et la contamination par des pesticides dangereux pour le corps humain [37], diagnostiquer les maladies des plantes et identifier les parasites [38]. En général, les applications du nez électroniques dans ce domaine sont destinées à remplacer les analyses de laboratoire longues et compliquées par des analyses rapides et simples qui se font sur le terrain.

Dans le domaine de l'environnement, en surveillant la qualité de l'air [39], du sol [40] et de l'eau [41], en effet, il est utilisé pour détecter les pollutions [42], les déchets liquides et les déversements toxiques [43], les émissions d'odeurs désagréables [44], et de gaz toxiques dangereux [45].

Dans le domaine de la médecine, en prédisant et identifiant les agents pathogènes qui génèrent un changement de l'odeur du patient qui peut être particulièrement perceptible dans l'haleine, dans l'urine ou dans les selles. Ces changements sont utilisés pour aider à

diagnostiquer la maladie, car l'odeur ou les fluides corporels du patient sont un outil de diagnostic important. Les analyses de l'odeur de l'haleine, l'urine ou les selles par le nez électronique peuvent être utilisées pour identifier et diagnostiquer des troubles tels que la dermatophytose [46], la pneumonie associée à la ventilation [47], le cancer [48], l'analyse rénale [49] et l'identification de l'asthme [50] ...etc.

Dans le domaine pharmaceutique, par le contrôle de la qualité des médicaments [51], et dans leurs formulations pharmaceutiques par l'analyse qualitative et quantitative de la cohérence des arômes ajoutées aux médicaments pour couvrir l'amertume [52].

Dans le domaine de la cosmétique, en développant des parfums pour valoriser le produit et augmenter l'attractivité des consommateurs en plus de la vérification de l'authenticité des produits haut de gamme [53].

Et dans le domaine de la sécurité en détectant les explosifs [54] et les matières inflammables [55]. Ainsi que dans de nombreux autres domaines.

IV. 4.7. Quelques nez électroniques commercialisés

Le tableau 1-3 répertorie quelques nez électroniques disponibles actuellement sur le marché. Les types de capteurs et les domaines d'application ou d'utilisation de ces nez électroniques commercialisés varient selon le cas.

Tableau 1-3 Nez électroniques commercialisés et leurs domaines d'utilisation [56]

Société	Produits	Domaine d'utilisation
Electronic sensor technology	Computer integrated zNose Battery operated zNose Portable zNose Benchtop zNose	Détection et analyse de composés organiques, biologiques et chimiques
AirSense analytics	Portable electronic nose (PEN)	Aéronautique, militaire, logistique, Sécurité publique, environnement,

		Détection de produits chimiques dangereux et d'explosifs
Electronic Nose Co. LTD.	Electronic Nose	Environnement, soins de santé, aliments et boissons
Sensigent	Cyranose eNose Aqua eNose QA	Surveillance des soins de santé et surveillance de l'air intérieur et extérieur
FOODsniffer	FOODsniffer	Alimentation (fraîcheur de la viande crue, de la volaille et du poisson)
EN Pty Ltd	E-Nose Mk3.3 E-Nose Mk3.3.0 E-Nose Mk4	Environnement, santé, sécurité
Alpha MOS	Heracles Neo	Aliments et boissons, emballage
Aryballe	NeOse Pro Aryballe EN	Automobile, soins personnels et cosmétiques, Appareils de consommation, aliments et boissons
Odotech	eNose SulfNose MultiNose	Environnement.
FIGARO Engineering Inc	Variété de capteurs de gaz et de modules de capteurs	Surveillance domestique et industrielle

IV. 4.8. Comparaison entre le nez humain et le nez électronique

- Le nez humain est doté de 10 millions de récepteurs qui se régénèrent automatiquement (naturellement) tandis que les nez électroniques sont composés de 3 à 100 capteurs chimiques qui se remplacent manuellement.
- Le nez humain identifie un grand nombre d'odeurs tandis que les nez électroniques doivent être formés pour chaque application.
- Le nez humain peut être infecté tandis que les nez électroniques peuvent être s'empoisonné.
- Les nez électroniques sont persistants tandis que le nez humain peut être saturé.
- Les nez électroniques peuvent détecter des molécules simples (H₂, H₂O, CO₂...) tandis que le nez humain ne peut pas détecter certaines d'entre elles.

V. QUELQUES METHODES ANALYTIQUES D'ANALYSE DES ODEURS

Les méthodes des analyses physico-chimiques sont des méthodes qui ont comme but la reconnaissance des composants d'un mélange odorant à travers l'utilisation des techniques analytiques classiques. Parmi les méthodes les plus utilisées on trouve celles qui suivent :

V.1. Chromatographie en phase gazeuse

Dans les années cinquante du siècle précédent plus précisément 1952, A.T. James et A.J.P. Martin ont inventé la chromatographie en phase gazeuse [57]. Et depuis son invention elle a montré son efficacité et son utilité dans divers domaine notamment dans le domaine des produits alimentaires en identifiant et en estimant plusieurs de ses composants. Le mot chromatographie vient des mots grecs chroma = couleur, et graphein = écrire, qui ensemble signifient littéralement « écrire en couleur ». La chromatographie en phase gazeuse est un outil d'analyse populaire et puissant, mais son majeur inconvénient réside dans la longueur du temps d'analyse.

La chromatographie sépare les mélanges en divisant les composants entre une phase mobile et une phase stationnaire. Cette technique est nommée en fonction de l'état physique de la phase stationnaire. Elle peut être appelée chromatographie gaz-solide si la phase stationnaire est solide, dans ce cas, la séparation est obtenue lorsque les composants ont des capacités d'adsorption différentes par rapport à la phase solide fixe. Ou chromatographie gaz-liquide si la phase stationnaire est liquide, dans ce cas, la séparation a lieu si les analytes ont des distributions

différentes entre les phases mobile et stationnaire. La température et la pression sont des facteurs importants qui peuvent affecter les résultats.

La figure 1-6 montre le schéma de principe de chromatographie en phase gazeuse. Après l'injection de l'échantillon dans le système, les molécules du mélange sont séparées au fur et à mesure que l'échantillon parcourt la longueur de la colonne - la partie principale du chromatographe placée dans une enceinte chauffée - en fonction de la différence des propriétés chimiques entre les différentes molécules et de leur affinité relative pour la phase stationnaire de la colonne.

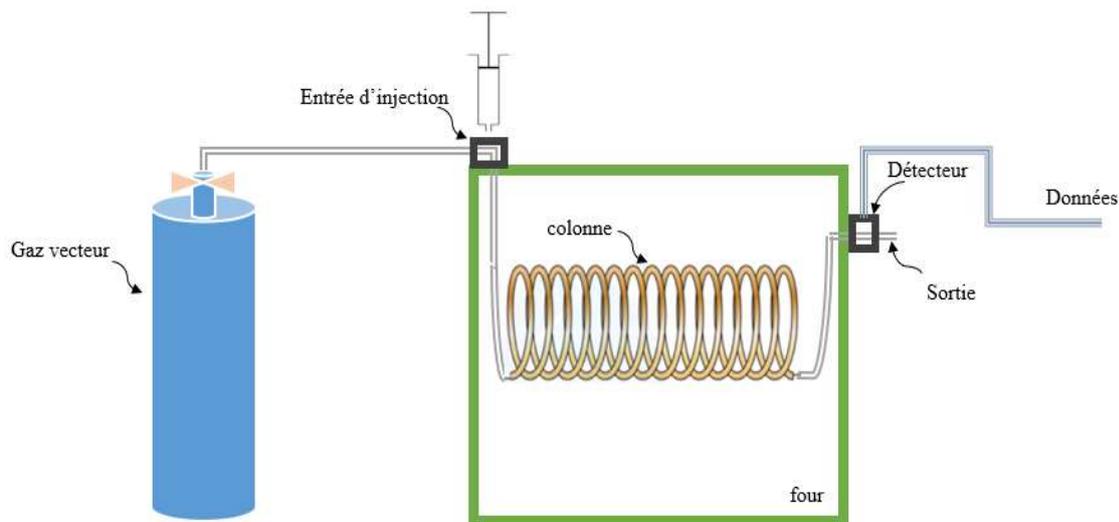


Figure 1-6 Schéma de principe d'un chromatographe en phase gazeux

V.2. Spectrométrie de masse

La spectrométrie de masse est apparue au début du vingtième siècle, et jusqu'à présent cette technique a connu un remarquable progrès [58]. Cette technique est une méthode d'analyse qui permet la détermination de la masse moléculaire des composés analysés, ainsi que leur identification et leur quantification. Son principe repose sur le fait que la composition des matériaux peut être connue en les exposant au champ électromagnétique, ce qui permet la séparation et la détection de leurs particules chargées en fonction de la différence de masse des atomes, molécules ou fragments moléculaires qui les composent. La figure 1-7 représente schématiquement un appareil de la spectrométrie de masse.

Après la première étape de l'ionisation où les atomes se font tomber un ou plusieurs électrons pour donner un ion positif, vient la deuxième étape de l'accélération pour que les atomes ionisés aient tous la même énergie cinétique. Ensuite les ions sont déviés par un champ

magnétique en fonction de leurs masses dans une troisième étape. La quantité de déviation dépend de leurs masses et du nombre de charges positives sur l'ion - en d'autres termes, du nombre d'électrons éliminés au cours de la première étape. Et enfin la détection électrique du faisceau d'ions dans la dernière étape.

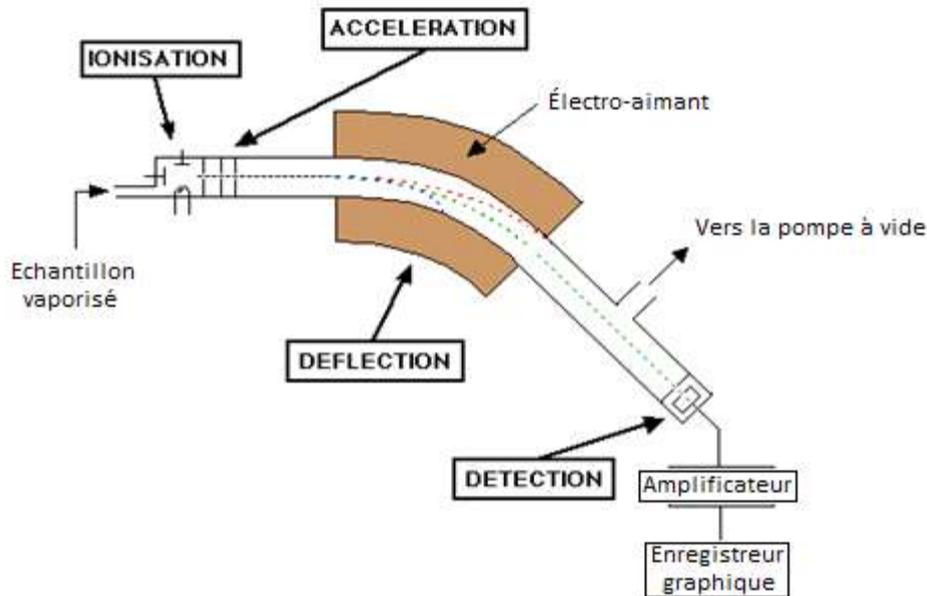


Figure 1-7 Principe d'un spectromètre de masse [59]

Malgré sa précision et son utilité, cette technologie est très coûteuse en termes d'équipement et d'entretien en plus de nécessiter des techniciens hautement spécialisés pour la manipuler.

V.3. Chromatographie en phase gazeuse couplée à la spectrométrie de masse

Comme son nom l'indique, la chromatographie en phase gazeuse couplée à la spectrométrie de masse est une technique qui rassemble entre la chromatographie en phase gazeuse et la spectrométrie de masse (figure 1-8), deux techniques déjà expliquées précédemment. Après l'injection de l'échantillon dans le système chromatographie en phase gazeuse, les molécules du mélange sont séparées au fur et à mesure que l'échantillon parcourt la longueur de la colonne en fonction de la différence des propriétés chimiques entre les différentes molécules et de leur affinité relative pour la phase stationnaire de la colonne. Les molécules sont ensuite introduites en spectrométrie de masse pour capturer, ioniser, accélérer,

dévier et détecter les molécules ionisées séparément. La spectrométrie de masse caractérise les molécules ionisées et fragmentées par leur rapport masse/charge.

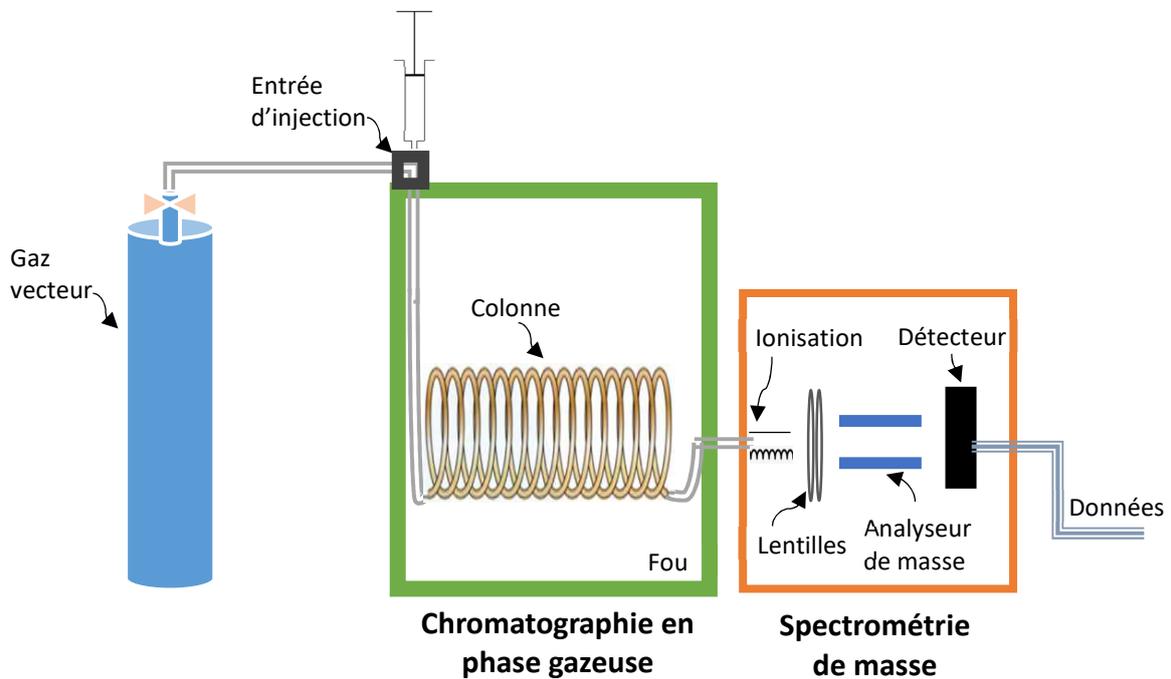


Figure 1-8 Chromatographe en phase gazeuse couplée à un spectromètre

V.4. Chromatographie en phase gazeuse couplée à l'Olfactométrie

La chromatographie en phase gazeuse couplée à l'Olfactométrie (figure 1-9) est une méthode où la machine et l'être humain se combinent pour la sélection des composants olfactifs d'un mélange complexe. Cette méthode publiée pour la première fois en 1964 [61] fournit à la

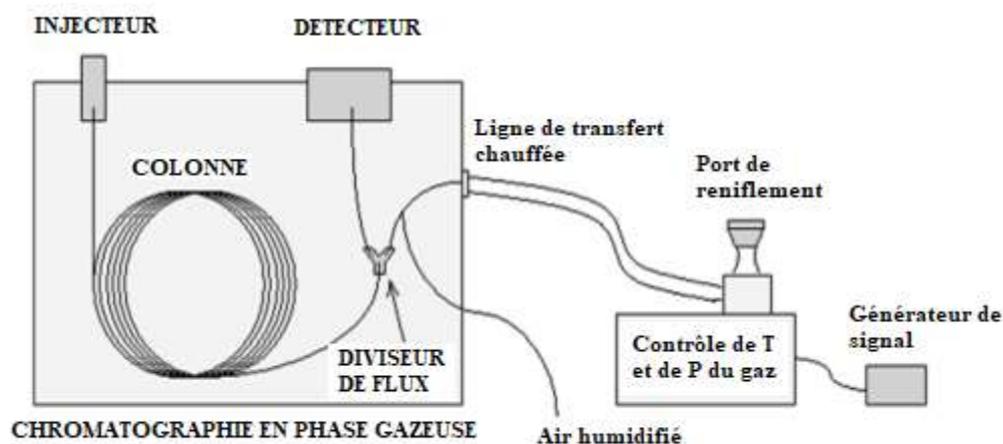


Figure 1-9 Chromatographie en phase gazeuse couplée à l'Olfactométrie [60]

fois une analyse instrumentale et sensorielle en utilisant le nez humain et est donc la seule technique analytique qui associe le pouvoir de résolution de la chromatographie en phase gazeuse capillaire à la sélectivité et à la sensibilité du nez humain.

Généralement dans cette méthode, la détermination de l'odeur de l'analyte se fait suivant deux moyens : grâce à un port dit olfactométrique dans lequel des personnes entraînées et sélectionnées pour leurs capacités olfactives sont utilisées, en parallèle à des détecteurs conventionnels, tels que le détecteur à ionisation de flamme ou le spectromètre de masse. L'éluat est séparé de telle sorte que les analytes atteignent les deux détecteurs simultanément, ce qui permet de comparer les deux signaux.

VI. CONCLUSION

Dans ce premier chapitre après une présentation de l'odeur et de ses caractéristiques ainsi que le système olfactif humain, nous avons présenté la technique du nez électronique qui se base sur l'utilisation d'une matrice des capteurs de gaz : un aperçu a été fait sur les différents types de capteurs avec une focalisation sur les capteurs de gaz à oxyde métallique, leur historique, leur structure interne, le principe de fonctionnement, caractéristiques et facteurs influents sur le bon fonctionnement de ces composants électroniques.

Les techniques physico-chimiques les plus utilisés pour l'analyse des mélanges de gaz qui peuvent être utilisés seuls ou couplés comme la chromatographie en phase gazeuse et la spectrométrie de masse (CG-SM) ont été présentés également. Ces méthodes sont efficaces pour les analyses mais elles souffrent de nombreux inconvénients qui limitent leurs utilisations à savoir le coût cher, le matériel lourd nécessitant un déplacement au laboratoire et une personne de spécialité pour la mise en œuvre...etc.

Références

- [1] M. Brancher, K. D. Griffiths, D. Franco, and H. de Melo Lisboa, "A review of odour impact criteria in selected countries around the world," *Chemosphere*, vol. 168, pp. 1531–1570, 2017, doi: 10.1016/j.chemosphere.2016.11.160.
- [2] C. Bushdid, M. O. Magnasco, L. B. Vosshall, A. Keller, and M. Mixture, "1 Trillion Olfactory Stimuli," *Science* (80-.), vol. 343, no. August, pp. 1370–1373, 2016.
- [3] S. Angeloni, A. M. Mustafa, D. Abouelenein, L. Alessandrini, L. Acquaticci, F. K. Nzekoue, R. Petrelli, G. Sagratini, S. Vittori, E. Torregiani and G. Caprioli "Characterization of the aroma profile and main key odorants of espresso coffee," *Molecules*, vol. 26, no. 13, pp. 1–29, 2021, doi: 10.3390/molecules26133856.

- [4] R. M. Fisher, R. J. Barczak, I. H. M. Suffet, J. E. Hayes, and R. M. Stuetz, "Framework for the use of odour wheels to manage odours throughout wastewater biosolids processing," *Sci. Total Environ.*, vol. 634, pp. 214–223, 2018, doi: 10.1016/j.scitotenv.2018.03.352.
- [5] J. M. Pinto, "Olfaction," *Proc. Am. Thorac. Soc.*, vol. 8, no. 1, pp. 46–52, 2011, doi: 10.1513/pats.201005-035RN.
- [6] Y. Soudry, C. Lemogne, D. Malinvaud, S. M. Consoli, and P. Bonfils, "Olfactory system and emotion: Common substrates," *Eur. Ann. Otorhinolaryngol. Head Neck Dis.*, vol. 128, no. 1, pp. 18–23, 2011, doi: 10.1016/j.anorl.2010.09.007.
- [7] L. B. Buc, "Smell and Taste: The Chemical Senses Odors Are Detected by Nasal Olfactory Sensory Neurons," pp. 1–30, 2010.
- [8] S. Nagayama, R. Homma, and F. Imamura, "Neuronal organization of olfactory bulb circuits," *Front. Neural Circuits*, vol. 8, pp. 1–19, 2014, doi: 10.3389/fncir.2014.00098.
- [9] D. A. Wilson, "Sensory Cortex Review," *Neurobiology*, vol. 11, pp. 498–504, 2001.
- [10] L. A. Camuñas-Mesa, B. Linares-Barranco, and T. Serrano-Gotarredona, "Neuromorphic spiking neural networks and their memristor-CMOS hardware implementations," *Materials (Basel)*, vol. 12, no. 7, 2019, doi: 10.3390/ma12172745.
- [11] A. Jinks and D. G. Laing, "A limit in the processing of components in odour mixtures," *Perception*, vol. 28, no. 3, pp. 395–404, 1999, doi: 10.1068/p2898.
- [12] F. Röck, N. Barsan, and U. Weimar, "Electronic nose: Current status and future trends," *Chem. Rev.*, vol. 108, no. 2, pp. 705–725, 2008, doi: 10.1021/cr068121q.
- [13] M. Jamal, M. R. Khan, S. A. Imam, and A. Jamal, "Artificial neural network based E-nose and their analytical applications in various field," *11th Int. Conf. Control. Autom. Robot. Vision, ICARCV 2010*, no. December, pp. 691–698, 2010, doi: 10.1109/ICARCV.2010.5707409.
- [14] R. W. Moncrieff, "An instrument for measuring and classifying odors," *Am. Physiol. Soc.*, 1961.
- [15] T. Seiyama, K. Fujiishi, M. Nagatani, and A. Kato, "A New Detector for Gaseous Components Using Zinc Oxide Thin Films," *J. Soc. Chem. Ind. Japan*, vol. 66, no. 5, pp. 652–655, 1963, doi: 10.1246/nikkashi1898.66.5_652.
- [16] W. F. Wilkens and J. D. Hartman, "an Electronic Analog for the Olfactory Processes," *Ann. N. Y. Acad. Sci.*, vol. 116, no. 2, pp. 608–612, 1964, doi: 10.1111/j.1749-6632.1964.tb45092.x.
- [17] A. Dravnieks and P. J. Trotter, "Polar vapour detector based on thermal modulation of contact potential," *J. Sci. Instrum.*, vol. 42, no. 8, pp. 624–627, 1965, doi: 10.1088/0950-7671/42/8/335.
- [18] P. J. Shaver, "Activated tungsten oxide gas detectors," *Appl. Phys. Lett.*, vol. 11, no. 8, pp. 255–257, 1967, doi: 10.1063/1.1755123.

- [19] N. Taguchi, A. E. Myer, and A. E. Geoffrey, "Gas Detective Device," US Pat., no. 54, pp. 1–5, 1971.
- [20] K. Persaud and G. Dodd, "Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose," *Nature*, vol. 299, no. 5881, pp. 352–355, 1982, doi: 10.1038/299352a0.
- [21] T. Oishi, M. Kaneyasu, and A. Ikegami, "Discrimination of Chemical Compounds and Functional Groups by Pattern Recognition Using an Integrated Sensor," *Microelectron. Int. An Int. J.*, vol. 5, no. 2, pp. 19–22, 1988, doi: 10.1108/eb044319.
- [22] J. W. Gardner and P. N. Bartlett, "A Brief History of ... Electronic Materials A Brief History of ... Electronic Materials," *Sensors Actuators B Chem.*, vol. 18, no. 19, pp. 211–220, 1994.
- [23] I. G. Sar van der, I. G. van der Sar, N. Wijbenga, G. Nakshbandi, J. G. J. V. Aerts, O. C. Manintveld, M. S. Wijsenbeek, M. E. Hellemons and C. C. Moor, "The smell of lung disease: a review of the current status of electronic nose technology," *Respir. Res.*, vol. 22, no. 246, pp. 1–22, 2021, doi.org/10.1186/s12931-021-01835-4.
- [24] J. Bardeen and W. H. Brattain, "Transistor, a semiconductor triode," *Proc. IEEE*, vol. 86, no. 1, pp. 29–30, 1998, doi: 10.1109/JPROC.1998.658753.
- [25] N. Taguchi, "Gas detecting device," United States Patent Office, Patented 3 Oct. 1972, pp. 1–7, 1972.
- [26] G. Heiland and H. Lamatsch, "Surface states on clean silicon," *Surf. Sci.*, vol. 2, no. C, pp. 18–25, 1964, doi: 10.1016/0039-6028(64)90038-X.
- [27] R. R. Ford and J. Pritchard, "Chemisorption of oxygen on gold," *Chem. Commun.*, no. 7, pp. 362–363, 1968, doi: 10.1039/C19680000362.
- [28] T. Seiyama and S. Kagawa, "Study on a Detector for Gaseous Components Using Semiconductive Thin Films," *Anal. Chem.*, vol. 38, no. 8, pp. 1069–1073, 1966, doi: 10.1021/ac60240a031.
- [29] G. M. Kale and J. M. Jafferson, "Solid-State Gas Sensors: Design and Fabrication," *Ref. Modul. Mater. Sci. Mater. Eng.*, no. 1, 2016, doi: 10.1016/b978-0-12-803581-8.01901-9.
- [30] J. Y. Kim, S. W. Kang, T. Z. Shin, M. K. Yang, and K. S. Lee, "Design of a Smart Gas Sensor System for Room Air- Cleaner of Automobile -Thick-Film Metal Oxide Semiconductor Gas Sensor-," *IEEE*, pp. 72–75, 2006.
- [31] J. K. Radhakrishnan, M. Kumara, and Geetika, "Effect of temperature modulation, on the gas sensing characteristics of ZnO nanostructures, for gases O₂, CO and CO₂," *Sensors Int.*, vol. 2, no. November 2020, p. 100059, 2021, doi: 10.1016/j.sintl.2020.100059.
- [32] Q. Qi, T. Zhang, X. Zheng, H. Fan, L. Liu, R. Wang, and Y. Zeng, "Electrical response of Sm₂O₃-doped SnO₂ to C₂H₂ and effect of humidity interference," *Sensors Actuators, B Chem.*, vol. 134, no. 1, pp. 36–42, 2008, doi: 10.1016/j.snb.2008.04.011.

- [33] A. Gliszczyńska-Świgło and J. Chmielewski, "Electronic Nose as a Tool for Monitoring the Authenticity of Food. A Review," *Food Anal. Methods*, vol. 10, no. 6, pp. 1800–1816, 2017, doi: 10.1007/s12161-016-0739-4.
- [34] A. Zakaria, A. Y. Md Shakaff, M. J. Masnan, M. N. Ahmad, A. H. Adom, M. N. Jaafar, S. A. Ghani, A. H. Abdullah, A. H. Abdul Aziz, L. M. Kamarudin, N. Subari and N. Ahmad Fikri, "A biomimetic sensor for the classification of honeys of different floral origin and the detection of adulteration," *Sensors*, vol. 11, no. 8, pp. 7799–7822, 2011, doi: 10.3390/s110807799.
- [35] C. Q. Wei, W. Y. Liu, W. P. Xi, D. Cao, H. J. Zhang, M. Ding, L. Chen, Y. Y. Xu and K. X. Huang, "Comparison of volatile compounds of hot-pressed, cold-pressed and solvent-extracted flaxseed oils analyzed by SPME-GC/MS combined with electronic nose: Major volatiles can be used as markers to distinguish differently processed oils," *Eur. J. Lipid Sci. Technol.*, vol. 117, no. 3, pp. 320–330, 2015, doi: 10.1002/ejlt.201400244.
- [36] C. Cevoli, L. Cerretani, A. Gori, M. F. Caboni, T. Gallina Toschi, and A. Fabbri, "Classification of Pecorino cheeses using electronic nose combined with artificial neural network and comparison with GC-MS analysis of volatile compounds," *Food Chem.*, vol. 129, no. 3, pp. 1315–1319, 2011, doi: 10.1016/j.foodchem.2011.05.126.
- [37] A. D. Wilson, "Applications of electronic-nose technologies for noninvasive early detection of plant, animal and human diseases," *Chemosensors*, vol. 6, no. 4, pp. 1–36, 2018, doi: 10.3390/chemosensors6040045.
- [38] S. Cui, P. Ling, H. Zhu, and H. M. Keener, "Plant pest detection using an artificial nose system: A review," *Sensors (Switzerland)*, vol. 18, no. 2, pp. 1–18, 2018, doi: 10.3390/s18020378.
- [39] G. Jasinski, L. Wozniak, P. Kalinowski, and P. Jasinski, "Evaluation of the Electronic Nose Used for Monitoring Environmental Pollution," 2018 15th Int. Sci. Conf. Optoelectron. Electron. Sensors, COE 2018, no. 22, pp. 1–4, 2018, doi: 10.1109/COE.2018.8435146.
- [40] U. Dorji, T. Pobkrut, and T. Kerdcharoen, "Electronic nose based wireless sensor network for soil monitoring in precision farming system," 2017 9th Int. Conf. Knowl. Smart Technol. Crunching Inf. Everything, KST 2017, pp. 182–186, 2017, doi: 10.1109/KST.2017.7886087.
- [41] J. W. Gardner, H. W. Shin, E. L. Hines, and C. S. Dow, "Electronic nose system for monitoring the quality of potable water," *Sensors Actuators, B Chem.*, vol. 69, no. 3, pp. 336–341, 2000, doi: 10.1016/S0925-4005(00)00482-2.
- [42] M. Iman, "Electronic noses for monitoring environmental pollution and building regression model," *IECON Proc. (Industrial Electron. Conf.)*, pp. 1730–1735, 2008, doi: 10.1109/IECON.2008.4758215.
- [43] A. D. Wilson, "Review of Electronic-nose Technologies and Algorithms to Detect Hazardous Chemicals in the Environment," *Procedia Technol.*, vol. 1, pp. 453–463, 2012, doi: 10.1016/j.protcy.2012.02.101.

- [44] C. Wongchoosuk, P. Lorwongtragool, and T. Kerdcharoe, "Malodor Detection Based on Electronic Nose," *Air Qual. Monit. Assess. Manag.*, 2011, doi: 10.5772/16582.
- [45] V. A. Ravindra and S. M. Rajbhoj, "Identification of Toxic Gases using Electronic Nose," 2017 Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2017, pp. 7–11, 2018, doi: 10.1109/ICCUBEA.2017.8463908.
- [46] N. Sahgal and N. Magan, "Fungal volatile fingerprints: Discrimination between dermatophyte species and strains by means of an electronic nose," *Sensors Actuators, B Chem.*, vol. 131, no. 1, pp. 117–120, 2008, doi: 10.1016/j.snb.2007.12.019.
- [47] C. Y. Chen, W. C. Lin, and H. Y. Yang, "Diagnosis of ventilator-associated pneumonia using electronic nose sensor array signals: Solutions to improve the application of machine learning in respiratory research," *Respir. Res.*, vol. 21, no. 1, pp. 1–12, 2020, doi: 10.1186/s12931-020-1285-6.
- [48] C. Baldini, L. Billeci, F. Sansone, R. Conte, C. Domenici, and A. Tonacci, "Electronic nose as a novel method for diagnosing cancer: A systematic review," *Biosensors*, vol. 10, no. 8, pp. 1–21, 2020, doi: 10.3390/bios10080084.
- [49] A. Voss, V. Baier, R. Reisch, K. V. Roda, P. Elsner, H. Ahlers and G. Stein, "Smelling renal dysfunction via electronic nose," *Ann. Biomed. Eng.*, vol. 33, no. 5, pp. 656–660, 2005, doi: 10.1007/s10439-005-1438-2.
- [50] P. Brinkman et al., "Identification and prospective stability of electronic nose (eNose)-derived inflammatory phenotypes in patients with severe asthma," *J. Allergy Clin. Immunol.*, vol. 143, no. 5, pp. 1811-1820.e7, 2019, doi: 10.1016/j.jaci.2018.10.058.
- [51] T. Wasilewski, D. Migoń, J. Gębicki, and W. Kamysz, "Critical review of electronic nose and tongue instruments prospects in pharmaceutical analysis," *Anal. Chim. Acta*, vol. 1077, pp. 14–29, 2019, doi: 10.1016/j.aca.2019.05.024.
- [52] L. Zhu, R. A. Seburg, E. Tsai, S. Puech, and J. C. Mifsud, "Flavor analysis in a pharmaceutical oral solution formulation using an electronic-nose," *J. Pharm. Biomed. Anal.*, vol. 34, no. 3, pp. 453–461, 2004, doi: 10.1016/S0731-7085(03)00651-4.
- [53] J. Gebicki, B. Szulczynski, and M. Kaminski, "Determination of authenticity of brand perfume using electronic nose prototypes," *Meas. Sci. Technol.*, vol. 26, no. 12, p. 125103, 2015, doi: 10.1088/0957-0233/26/12/125103.
- [54] J. W. Gardner, "Review of Conventional Electronic Noses and Their Possible Application to the Detection of Explosives," *Electron. Noses Sensors Detect. Explos.*, pp. 1–28, 2004, doi: 10.1007/978-1-4020-2800-7_1.
- [55] Z. Wu, H. Wang, X. Wang, H. Zheng, Z. Chen, and C. Meng, "Development of electronic nose for qualitative and quantitative monitoring of volatile flammable liquids," *Sensors (Switzerland)*, vol. 20, no. 7, 2020, doi: 10.3390/s20071817.
- [56] D. Karakaya, O. Ulucan, and M. Turkan, "Electronic Nose and Its Applications: A Survey," *Int. J. Autom. Comput.*, vol. 17, no. 2, pp. 179–209, 2020, doi: 10.1007/s11633-019-1212-9.

- [57] T. Feng, M. Sun, S. Song, H. Zhuang, and L. Yao, Gas chromatography for food quality evaluation. Elsevier Inc., 2019.
- [58] X. Zhang, Mass spectrometry for food quality and safety evaluation, no. 5. Elsevier Inc., 2019.
- [59] J. Clark, “How the spectrometer works,” LibreTexts, pp. 1–3, 2021, doi: <https://chem.libretexts.org/@go/page/3725>.
- [60] B. Plutowska and W. Wardencki, “Application of gas chromatography-olfactometry (GC-O) in analysis and quality assessment of alcoholic beverages - A review,” Food Chem., vol. 107, no. 1, pp. 449–463, 2008, doi: 10.1016/j.foodchem.2007.08.058.
- [61] G. H. Fuller, R. Steltenkamp, and G. A. Tisserand, “the Gas Chromatograph With Human Sensor: Perfumer Model,” Ann. N. Y. Acad. Sci., vol. 116, no. 2, pp. 711–724, 1964, doi: 10.1111/j.1749-6632.1964.tb45106.x.

Chapitre 2 : ANALYSE DE DONNEES ET APPRENTISSAGE AUTOMATIQUE

I. INTRODUCTION

Ce chapitre décrira les différentes techniques de classification et régression de données utilisées conjointement avec les dispositifs multi-capteurs de gaz (nez électronique). Mais avant d'arriver à l'utilisation de ces algorithmes, il est indispensable de passer par un prétraitement de données. Dans ce chapitre, nous présenterons les techniques utilisées afin de préparer les données brutes générées par les capteurs de gaz, et quelques techniques de validation des modèles.

La rédaction de ce chapitre sera organisée comme suivant :

- Dans la première partie nous allons présenter le prétraitement de données :
 - Le centrage et la mise à l'échelle ;
 - L'extraction des caractéristiques (Features extraction) ;
 - La normalisation.

- Dans la seconde partie, nous présenterons quelques méthodes d'analyses multivariées de données en commençant par :
 - Quelques méthodes de la classification comme :
 - L'Analyse en Composante Principale,
 - Les Séparateurs à Vastes Marges,
 - Les K plus Proches Voisins.

 - Quelques méthodes de la régression comme :
 - Régression par moindres carrés partielles,
 - Régression des Séparateurs à Vaste Marge,
 - Réseau Autorégressif non linéaire à entrées exogènes.

- Dans la troisième et dernière partie, nous présenterons aussi quelques techniques de validation des modèles :
 - Hold-out Validation,
 - K-Fold cross-validation.
 - Leave-one-out cross-validation.

II. PRETRAITEMENT DE DONNEES

Le prétraitement des données est une manipulation de données dans le but de les préparer et les transformer sous la forme la plus appropriée pour l'analyse. C'est une étape très importante et nécessaire pour structurer les données car un prétraitement approprié peut créer une différence entre un modèle utile et inutile ainsi que l'amélioration des performances. Après la collecte des réponses brutes des capteurs, ces dernières sont converties de leurs formes analogiques en un signal numérique qui peut être interprété par un ordinateur. Ensuite, les données sont centrées et nettoyées en éliminant les données hors limites, en extrayant et en sélectionnant les caractéristiques clés et en éliminant l'effet d'échelle par normalisation.

II.1. Le centrage et la mise à l'échelle

Le centrage et la mise à l'échelle des données sont les procédures de transformation de données les plus fondamentales et les plus simples. Avant d'entrer dans le modèle, plusieurs approches d'exploration de données exigent que les données soient centrées et mises à l'échelle.

Le centrage est une technique de prétraitement « variable par variable » qui conduit à la suppression des décalages présumés. Chaque valeur d'une variable dans l'ensemble de données est soustraite de la valeur moyenne pour la centrer dans l'ensemble de données ; cela implique que la distribution des fluctuations autour de la moyenne de la variable est changée en celle autour de zéro. En conséquence, la propriété fluctuante de la variable est centrée, ne laissant que le changement entre les observations pour l'étude. De même, pour la mise à l'échelle d'une variable, toutes les valeurs de la variable sont divisées par l'écart type de cette dernière, et les variables associées sont placées sur un pied d'égalité en termes de variation. La mise à l'échelle n'est pas nécessaire si toutes les variables de l'ensemble de données sont mesurées par la même unité [1].

II.2. Extraction des caractéristiques (Features extraction)

L'extraction de caractéristiques est une étape très importante, elle sert à extraire et à trouver un nombre de caractéristiques potentiellement bonnes (le plus distinctif et informatif) pour former une base solide et arriver à un traitement performant des données. Un nombre élevé de paramètres peut être pénalisant en termes de temps de calcul.

Cette étape nécessite l'observation et l'examen des réponses temporelles des capteurs afin d'extraire les informations pertinentes, distinctives et informatives pour une classification

ou régression efficace. Les performances des algorithmes d'apprentissage automatique qui seront utilisés par la suite reposent sur l'efficacité des caractéristiques choisies, si les caractéristiques extraites ne désignent pas clairement les signaux employés et inappropriés, un algorithme de classification exploitant de telles caractéristiques pourrait conduire à des problèmes d'identification des classes des caractéristiques. Par conséquent, la précision de la classification peut être faible.

Généralement, l'extraction des caractéristiques se fait directement à partir des courbes de réponse temporelles des réseaux de capteurs, trois classes de paramètres peuvent être distinguées : les variables extrêmes, les variables issues du domaine transitoire de réponse et celles qui sont liées à la forme du signal.

Le tableau 2-1 suivant est le résumé de quelques caractéristiques les plus utilisées dans l'extraction de données :

Tableau 2-1 Caractéristiques les plus utilisées dans l'extraction des données pertinentes et leurs descriptions

Caractéristique	Description
Réponse maximale	La valeur maximale générée par le capteur
Réponses dans un temps spécial	La valeur de la réponse dans un temps spécial sur la courbe de réponse
Temps spécial d'une réponse	Le temps d'une réponse spéciale sur la courbe de réponse
Superficie (aire)	Les valeurs de surface des courbes du capteur et de l'axe des temps entourés
Intégrale	$I = \int_a^b x(t) dt$
Dérivée	$D' = \frac{dx(t)}{dt}$
Différence	$x(t_b) - x(t_a)$

II.3. Normalisation

Après la phase de l'extraction de données, les principales caractéristiques résultantes des signaux bruts d'un réseau de capteurs nécessitent une normalisation pour compenser les différences d'échelle afin qu'elles soient utilisées comme une entrée utile dans l'analyse des données et sans que des composantes n'agissent plus que les autres sur le résultat de la classification ou la régression.

La normalisation est la dernière étape du prétraitement numérique avant l'analyse multivariée. La diversité des méthodes présentes pour cette opération donne la possibilité du choix de la stratégie la plus adaptée selon le cas étudié, il n'y a pas de règle générale pour connaître la technique de pré-traitement des données appropriée au vu du cas étudié ou des appareils utilisés.

Le tableau 2-2 suivant présente quelques méthodes standards de normalisation [2] des données les plus utilisés pour une matrice de données X_{ij} :

Tableau 2-2 Méthodes les plus utilisées dans la normalisation des données

Méthode	Description
Soustraction de la ligne de base	$x_{ij} = x_{ij} - x_{il}$
Linéarisation	$x_{ij} = \sqrt{x_{ij}}$ Ou $x_{ij} = \log(x_{ij})$
Moyenne centrée	$x_{ij} = x_{ij} - \bar{x}_j$
Echelle de gamme	$x_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)}$
Mise à l'échelle relative	$x_{ij} = \frac{x_{ij}}{\max(X)}$ $x_{ij} = \frac{x_{ij}}{\max(x_i)}$
Mise à l'échelle automatique	$x_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_{X_j}}$

III. METHODES D'ANALYSES MULTIVARIABLES DE DONNEES

Les méthodes d'analyses multivariées représentent la phase finale des systèmes nez électroniques. L'analyse multivariée se compose d'un ensemble de techniques qui peuvent être utilisées lorsque plusieurs mesures sont prises sur chaque élément dans un ou plusieurs échantillons. Ces techniques sont utilisées pour créer des empreintes digitales basées sur des observations du fonctionnement du système en apprenant. Après avoir terminé cette tâche, le système sera prêt à identifier les éléments inconnus qu'il va rencontrer.

Dans l'analyse multivariée, il existe deux catégories de méthodes de base : l'apprentissage non supervisé et supervisé. On parle d'un apprentissage non supervisé lorsqu'il n'y a pas d'étiquettes pour les données et que seule la structure sous-jacente des données peut être utilisée pour apprendre comme la méthode d'analyse en composantes principales (ACP) et la méthode k-moyennes. Un apprentissage est dit supervisé lorsque les données sont entièrement étiquetées pour la formation du modèle, c'est-à-dire la pré-connaissance des classes des données. Dans l'apprentissage automatique de données, il existe deux catégories de sous-problèmes : la classification et la régression. On parle de la classification lorsque les valeurs cibles à prédire sont discrètes, parmi les algorithmes de classification les plus connus on trouve les séparateurs à vastes marges (SVM) et K-plus proches voisins (KNN). Et de la régression lorsque les valeurs cibles sont continues, comme la méthode de la régression des moindres carrés partiels (PLSR) et le réseau autorégressif non linéaire à entrées exogènes (NARX).

III.1. Méthodes de classification des données

III.1.1. Analyse en Composantes Principales (ACP)

L'analyse en composantes principales (ACP), créée par Karl Pearson en 1901 [3], intégrée aux statistiques mathématiques par Harold Hotelling en 1933 [4], est l'une des approches les plus fréquemment utilisées en traitement du signal pour réduire les dimensions des matrices multidimensionnelles à une base de deux ou trois coordonnées tout en minimisant la perte de l'information. C'est une méthode de reconnaissance de forme non supervisée qui sert comme outil visuel pour afficher graphiquement les regroupements et les séparations au sein des ensembles de données souvent utilisée pour explorer les données d'un réseau de capteurs de gaz [5]. Elle fonctionne en extrayant les informations les plus importantes d'une collection de données et en les représentant comme un ensemble de nouvelles variables orthogonales connues sous le nom de composantes principales (PC, en anglais : Principal Component).

III.1.1.1. Principe de l'ACP

L'objectif de l'ACP est l'extraction des informations appropriées contenues dans un tableau de données de grande dimension. Afin d'atteindre cet objectif, le tableau de données doit être réduit et l'information doit être compressée en supprimant la majorité de sa structure en gardant les informations essentielles pour créer des représentations graphiques à la fois fidèles, précises et simples à comprendre [6]. L'idée est de projeter les données dans un espace avec moins de dimensions, généralement deux ou trois, afin qu'elles puissent être lues.

Prenons par exemple le cas d'une matrice de données (N x D) avec N individus allant de 1, ... i, ..., N et D variables quantitatives notées $x_1, \dots, x_j, \dots, x_D$

$$\begin{array}{c} \text{Variables quantitatives} \\ \text{Individus} \end{array} \begin{bmatrix} x_{11} & \dots & x_{1j} & \dots & x_{1D} \\ \dots & & & & \\ x_{i1} & \dots & x_{ij} & \dots & x_{iD} \\ \dots & & & & \\ x_{N1} & \dots & x_{Nj} & \dots & x_{ND} \end{bmatrix}$$

L'ACP commence par la normalisation des N individus dans le tableau des données d'origine. La normalisation consiste à faire centrer et réduire le nuage des individus selon l'équation :

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_{x_{ij}}}$$

Où x'_{ij} est la matrice centrée réduite, x_{ij} est le tableau de données initial, \bar{x}_j est le point moyen pour la variable j et $\sigma_{x_{ij}}$ est l'écart type de cette variable.

La seconde étape de la méthode ACP consiste à obtenir les valeurs et vecteurs propres à partir de la matrice de corrélation R qui sert à évaluer la dépendance entre les variables. La matrice de corrélation R est calculée selon l'équation :

$$R = X^T \cdot X$$

L'ellipsoïde d'inertie dont on cherche les axes de symétrie est représenté par cette matrice. Une fois la matrice de corrélation est calculée, les coordonnées des points-individus sont alors déterminées sur de nouveaux axes factoriels dits axes principaux.

III.1.1.2. Analyse du nuage des points-individus et points variables

La procédure de détermination du plan dit principal consiste à repérer avant tout la droite $\Delta 1$ qui génère la plus grande moyenne des carrés des distances entre les projections des individus sur cette droite. Puis, en reliant deux à deux des axes orthogonaux entre eux, on trouve un axe $\Delta 2$ perpendiculaire à $\Delta 1$ qui donne la plus grande moyenne des carrés des distances entre les projections sur cet axe, et ainsi de suite. Les grands axes du nuage sont ainsi identifiés. Par conséquent, en projetant l'individu i avec les coordonnées de départ $x_{i1}, x_{i2}, \dots, x_{ip}$, les nouvelles coordonnées $c_{i1}, c_{i2}, \dots, c_{ip}$, seront fournies par l'équation ci-dessous :

$$CP^k = u_1^k x_1 + u_2^k x_2 + \dots + u_p^k x_p$$

Où $x_{i1}, x_{i2}, \dots, x_{ip}$ sont les caractères mesurés sur les n individus et les coefficients $u_1^k, u_2^k, \dots, u_p^k$ forment les $k^{\text{ième}}$ facteurs principaux.

En effet, ces composants primaires seront les vecteurs propres de la matrice de corrélation V , qui seront liés aux valeurs propres dérivées en diagonalisant la matrice V . Pour chaque axe principal k obtenu, est associé une valeur propre qui représente le pourcentage d'inertie ou encore la variance donnée par l'axe k . Sur les deux premiers axes, cette partie d'inertie représente l'aplatissement du nuage. Dans le cas général, une meilleure représentation du nuage de points dans ce plan factoriel est montrée par une grande partie d'inertie.

Contrairement à celui des points-individus, pour constituer le nuage des points-variables, l'analyse ne se fait pas par rapport au centre de gravité du nuage de points-variables dans l'espace \mathfrak{R}^n mais par rapport à l'origine. En effet, à partir de la matrice centrée-réduite X , on exprime la distance d'une variable j à l'origine par la relation suivante :

$$d^2(0, j) = \sum_{i=1}^n (x_{ij})^2 = 1$$

Tous les points-variables sont sur une sphère, de rayon 1 centrée à l'origine des axes, appelée sphère de corrélation. L'objectif est alors de définir les plans d'ajustement suivant des cercles de rayon 1 (cercles des corrélations) à l'intérieur desquels se trouveront les points-variables.

Les variables sont alors projetées sur les mêmes plans obtenus précédemment pour les points-individus. Cette fois, la coordonnée d'un point-variable sur un axe k sera le coefficient de corrélation de cette variable avec le k^{ième} facteur principal qui est une combinaison linéaire des variables initiales. Cette coordonnée est donc nécessairement inférieure à 1 en valeur absolue.

Grâce à l'ACP, on peut donc également mettre en évidence la proximité entre deux points-variables en termes de corrélation. Ainsi, un faible coefficient de corrélation entre deux variables signifie qu'elles sont indépendantes linéairement. Cependant, il peut exister des liaisons non linéaires qui ne sont pas mises en évidence par une telle méthode.

III.1.1.3. Algorithme de l'ACP

- 1 : Entrée : un ensemble d'apprentissage D-dimensionnel $X = \{x_1, x_2, \dots, x_n\}$
 - 2 : Calculer la moyenne $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$
 - 3 : Calculer la matrice de covariance $cov(x) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}_i)(x_i - \bar{x}_i)^T$
 - 4 : Trouver la décomposition spectrale de Cov(x), en obtenant les vecteurs propres $\xi_1, \xi_2, \dots, \xi_D$ et leurs valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_D$.
- Notez que les valeurs propres sont triées $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_D \geq 0$
- 5 : pour toute $x \in \mathfrak{R}^D$, sa nouvelle représentation dimensionnelle inférieure est :

$$y = \{\xi_1^T(x - \bar{x}), \xi_2^T(x - \bar{x}), \dots, \xi_d^T(x - \bar{x})\} \in \mathfrak{R}^d$$

III.1.2. Séparateurs à Vaste Marge (SVM)

Les machines à vecteurs de support (SVM, en anglais Support Vector Machines) sont les approches les plus résilientes et les plus précises parmi tous les algorithmes bien connus dans les applications d'apprentissage automatique d'aujourd'hui [7]. Cette technique initiée par Vapnick [8] en 1995 est fondée sur de solides bases théoriques.

La méthode des Séparateurs à Vaste Marge (SVM) est un algorithme d'apprentissage supervisé qui classe les données à l'aide de la fonction de noyau linéaire ou non linéaire, cette fonction tente de distinguer les classes par des hyper-plans. Les hyperplans sont trouvés tels qu'ils n'incluent pas de données intérieures et fournissent une marge maximale entre les paires de classes. Les points de données situés sur ces hyperplans sont appelés vecteurs de support. L'hyperplan passant par le centre de la distance maximale séparant les deux dits hyperplans est le meilleur résultat (figure 2-1).

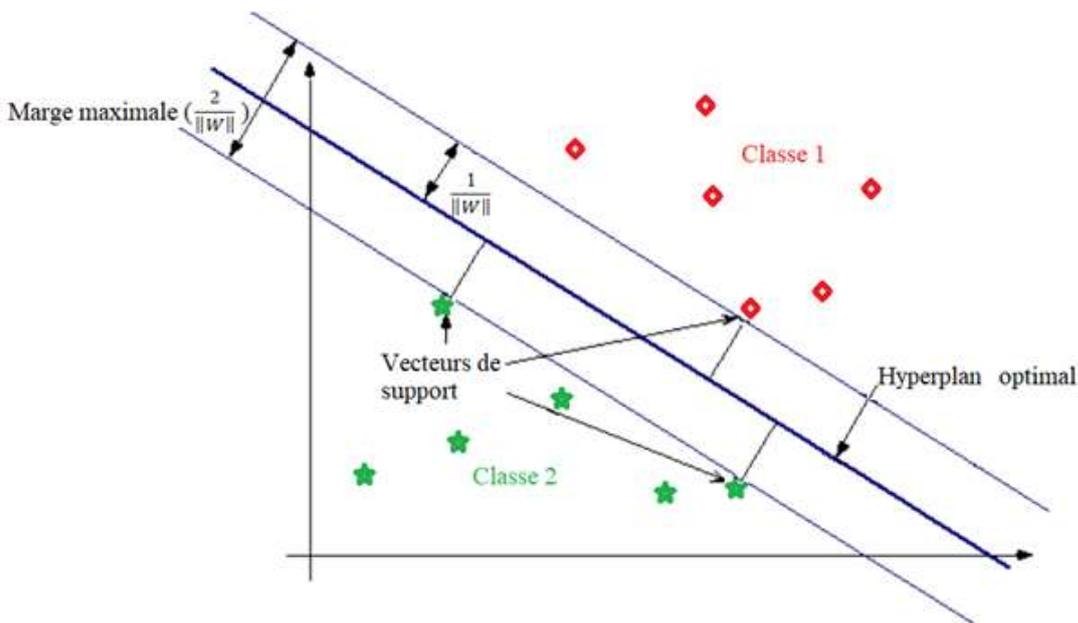


Figure 2-1 Exemple d'un hyperplan optimal dans \mathcal{R}^2

III.1.2.1. SVM avec classificateur linéaire

L'utilisation originale des algorithmes SVM était de déterminer la meilleure séparation entre deux classes [9]. Voilà un exemple de problème de classification à deux classes :

Soit X un ensemble de N exemples d'apprentissage dont chaque élément est représenté par une paire (\vec{x}_i, y_i) avec $i = 1 \cdot \cdot \cdot N$, y_i un label de classe pouvant prendre la valeur $+1$ ou -1 et \vec{x}_i un vecteur de dimension d ($\vec{x}_i \in \mathcal{R}^d$). L'objectif d'un classificateur est d'établir une fonction de décision $f : (x, \alpha) \rightarrow y$ (avec α paramètres du classificateur), en utilisant les données d'apprentissage, afin de l'exploiter postérieurement pour classer de nouvelles données.

Pour un classificateur linéaire, la fonction f peut être définie à l'aide d'un hyperplan d'équation $\vec{w} \cdot X + b = 0$ (où \vec{w} et b indiquent les paramètres de l'hyperplan soit successivement un vecteur normal au plan et le biais). La classification d'un vecteur \vec{x}_i est ainsi donnée par le signe de la fonction f , $\text{sing}[f_{\vec{w},b}(\vec{x}_i)]$ c-à-d que $y_i = +1$ si $\text{sing}[f_{\vec{w},b}(\vec{x}_i)] > 0$ ou $y_i = -1$ si $\text{sing}[f_{\vec{w},b}(\vec{x}_i)] < 0$. Le principe de la méthode SVM est de chercher le meilleur hyperplan permettant d'assigner convenablement les données avec une marge maximale.

III.1.2.2. SVM avec classificateur non-linéaire

Dans ce cas, l'objectif du classificateur SVM est de changer l'espace de données afin d'éviter les inconvénients des cas non linéairement séparables. Cette transformation offre l'occasion de distinguer linéairement les données dans un nouvel espace. En conséquence, nous allons avoir un changement de dimension. Le terme « espace de re-description » décrit cette nouvelle dimension. Certainement, plus la dimension de l'espace de re-description est large, plus il est susceptible de localiser un hyperplan séparateur entre les exemples. La figure 2-2 illustre cela.

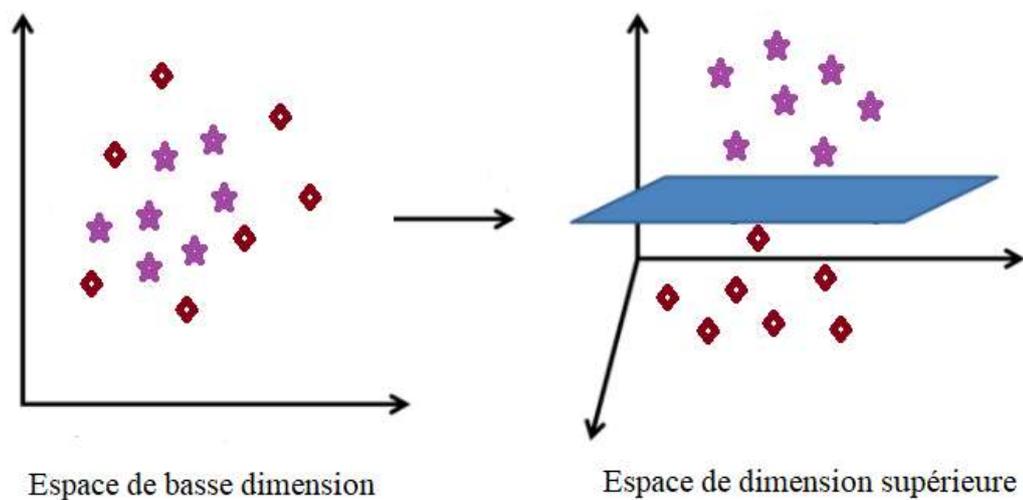


Figure 2-2 Représentation de l'astuce du changement de l'espace

En effet, nous pouvons convertir un problème de séparation non linéaire dans l'espace de représentation en un problème de séparation linéaire dans un espace de re-description de dimension supérieure. Une fonction noyau (kernel) introduite pour la première fois par Aizerman et col [10] est utilisée pour exécuter cette transformation non linéaire, elle transforme tout algorithme linéaire qui dépend uniquement de produits scalaires entre vecteurs en une contrepartie non linéaire. Pratiquement, seules quelques familles de fonctions paramétrables du noyau sont connues, et il appartient à l'utilisateur de SVM d'évaluer celle la plus convenable à son application. Les noyaux polynomiaux, gaussiens, radial, sigmoïdes et laplaciens sont quelques exemples de noyaux.

III.1.2.3. Classification multi-classes par SVM

Jusqu'à maintenant dans l'explication du SVM, nous n'avons abordé que ce qui peut être appliqué dans les cas de la classification binaire. Pour les données qui doivent être classées en

classes M , deux manières sont possibles. La première consiste à décomposer le problème initial en un certain nombre de sous-problèmes à deux classes, puis à utiliser les résultats de l'application de SVM standard à chacun de ces sous-problèmes pour réduire le problème à un problème binaire. La deuxième consiste à construire un modèle de classification multi-classes en généralisant SVM décrit auparavant [11].

Une technique pour réduire le problème à un problème de classification binaire consiste à classer les points d'une classe (les points marqués $+1$) par rapport aux points de toutes les autres classes (les points marqués -1). Cette technique est connue sous le nom "one-against-the-rest" (1-v-r) [12]. En conséquence, M frontières de séparation sont déterminées pour M classes. Les fonctions de décision acquises par classification binaire sont utilisées pour prévoir la classe d'un nouveau point : $y = \text{sign}(f(w_i, x) + b_i)$, $i = 1 \dots M$. Au lieu d'utiliser la fonction de signe, qui ne renvoie que -1 ou 1 , nous utiliserons la classe qui maximise $f(w_i, x) + b_i$. Cette méthode de procéder par réponse maximale est similaire à la classification multi-classes du Perceptron Multicouches (PMC).

La technique "1-v-r", en revanche, présente un sérieux défaut, c'est pourquoi elle n'est pas largement utilisée de nos jours. En réalité, tout entraînement individuel utilise deux ensembles de données asymétriques, si la quantité de données de formation dans chaque classe est presque la même. Le nombre des données avec l'étiquette -1 est toujours supérieure au nombre des données avec l'étiquette $+1$. De plus, chaque entraînement intègre toutes les données de la base ce qui nécessite un temps important de machine.

Une autre méthode consiste à évaluer toutes les paires de classes possibles et à générer un hyperplan pour chaque paire de classes en utilisant l'approche un contre un "one-against-one" (1-v-1) [13]. On obtient alors $\frac{M(M-1)}{2}$ hyperplans séparateurs. Chaque hyperplan vote pour sa classe choisie lorsqu'une nouvelle observation se produit, et la classe avec le plus de votes gagne [14]. D'autres procédures de prise de décision des hyperplans $\frac{M(M-1)}{2}$ ont également été proposées. Par rapport à la technique "1-v-r", nous pouvons observer que cette méthode évite le problème d'asymétrie dans l'ensemble d'apprentissage. Plus de formation semble être nécessaire dans ce scénario, mais comme seules deux classes sont prises en compte à chaque fois, nous pouvons toujours nous attendre à un gain de temps.

Pour les approches "1-v-r" et "1-v-1", la combinaison est réalisée respectivement par le maximum et un vote. L'approche code correcteur d'erreur ECOC est une autre approche

fréquemment utilisée dans le cadre du SVM (Error Correcting Output Coding) [15]. Dans cette approche, la classe d'un individu est décrite par un mot codé de N bits. En conséquence, nous avons une matrice de codes avec les classes en lignes et les classificateurs en colonnes. La contribution de chaque classificateur au résultat final de la classification est représentée par la matrice de codes, qui est basée sur les erreurs des divers classificateurs. Par conséquent, la classe d'un nouvel individu est déterminée par la classe dont le code est le plus proche en termes de distance.

III.1.2.4. Algorithme de SVM

L'algorithme de la méthode SVM pour deux classes est comme suit :

1. Normaliser les données $M = \frac{x-\mu}{\sigma}$ avec $\mu = \frac{\sum x}{n}$ et $\sigma = \sqrt{\frac{1}{n-1} \times \sum (x_i - \mu)^2}$
2. Calculer la matrice augmentée $[M - e]$ c'est-à-dire augmenter la matrice de colonne "-1" à M.
3. Calculer la matrice diagonale H avec les classes ici c'est 1 et -1 : $H = D[M - e]$ et $H^T H$
4. Calculer $U = V \times \left[I - H \left[\frac{I}{V} + H^T H \right]^{-1} H^T \right] \times e$ Où I = Matrice d'Identité et V = Ordre de $H^T H$ avec une valeur de 0,1.
4. Calculer $w = M^T D U$ et $\gamma = -e^T D U$
5. Calculer $w^T x - \gamma$
6. Comparer le signe de $w^T x - \gamma$ avec l'étiquette de classe d'entrée.

L'algorithme de la méthode SVM pour deux classes ou le classificateur binaire est la base de la méthode SVM. Un classificateur SVM multi-classe est composé de nombreux classificateurs SVM binaires combinés en un seul classificateur. Un vote majoritaire entre les classificateurs SVM binaires est utilisé pour classer dans un SVM multi-classe.

III.1.3. K plus Proches Voisins (KNN)

III.1.3.1. Présentation et principe de la méthode KNN

Proposée la première fois par Fix et Hodges [16] en 1951 et modifiée par Cover en 1967 [17], la méthode des k-plus proches voisins (KNN, en anglais : K - Nearest Neighbors) est une méthode simple et puissante qui repose sur l'affectation de l'échantillon d'entrée à la classe à laquelle appartiennent la majorité des k échantillons les plus proches de l'ensemble

d'apprentissage [18], cette méthode non paramétrique supervisée est capable de faire une classification non linéaire à partir d'un petit nombre d'échantillons.

Il y a trois éléments clés de cette approche : un ensemble d'objets étiquetés, une métrique de distance ou de similarité pour calculer la distance entre les objets, et la valeur de k , le nombre de voisins les plus proches. Pour classer un objet non étiqueté, la distance de cet objet aux objets étiquetés est calculée, ses k plus proches voisins sont identifiés, et les étiquettes de classe de ces plus proches voisins sont ensuite utilisées pour déterminer l'étiquette de classe de l'objet. Une fois la liste des voisins les plus proches est obtenue, l'objet de test est classé en fonction de la classe majoritaire de ses plus proches voisins.

Le choix des éléments clés affecte les performances de cette approche. L'un est le choix de k , le bon choix de K est très intéressant dans cet algorithme car il influence les résultats obtenus. Si k est trop petit, le résultat peut être sensible aux points de bruit tandis que la plupart du temps, des valeurs plus grandes de k réduisent la valeur de la classification en incluant trop de points d'autres classes.

La figure 2-3 illustre l'importance du bon choix du facteur k car en choisissant $k=3$ le nouvel exemple à classer serait attribué à la deuxième classe et si $k=7$ il serait attribué à la première.

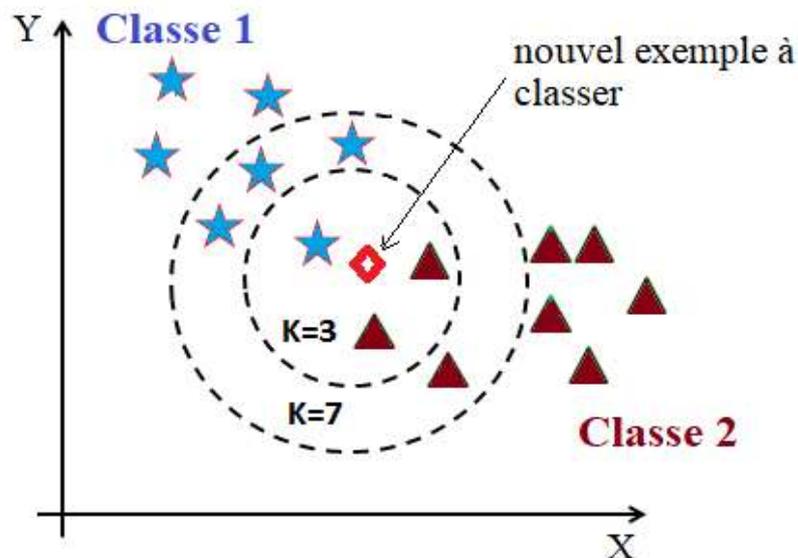


Figure 2-3 Problème du choix de facteur k pour la méthode KNN

Un autre facteur crucial est la mesure de la distance. Bien qu'une variété de mesures de distance puisse être utilisée pour calculer la distance entre deux points, la mesure de distance la

plus désirée est celle dans laquelle une distance moindre entre deux éléments indique une plus grande possibilité d'appartenir à la même classe. La grande dimensionnalité des données peut également avoir un impact sur certaines métriques de distance. Il est bien connu qu'au fur et à mesure que le nombre de qualités augmente, la métrique de distance euclidienne devient moins discriminante. De plus, il peut être nécessaire d'ajuster les caractéristiques pour éviter que l'une des qualités ne domine les mesures de distance. Malgré cela, la distance euclidienne est la métrique la plus utilisée, elle est calculée comme suit :

$$D = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

Il existe d'autres mesures de distance [19] telles que :

- La distance de Manhattan : $D = \sum_{i=1}^n |p_i - q_i|$
- La distance de Minkowski : $D = (\sum_{i=1}^n |p_i - q_i|^a)^{1/a}$

III.1.3.2. Algorithme de KNN

Entrées :

$x \in D$, avec D un ensemble de données ;

X un ensemble de données étiquetées ;

$K \in \{1, \dots, N\}$;

d une mesure de similarité (distance) définie sur D .

Sorties :

La classe de x .

1 : **pour chaque** $x_i \in X$ faire

2 : calculer $d(x_i, x)$

3 : **Fin pour**

4 : **pour chaque** $k \in \{1, \dots, k\}$ faire

5 : $KNN[k] \leftarrow \text{argmind}(x_i, x)$

($x_i \in X$)

6 : $d(x_i, x) \leftarrow +\infty$

7 : **Fin pour**

8 : **Attribuer** x à la classe la plus fréquente (ou la classe majoritaire)

III.2. Indicateurs d'évaluation des méthodes de classification

En plus des figures représentatives générées pour chaque algorithme, un résumé des résultats de prédiction est utilisé pour évaluer les performances des modèles exploités, c'est la matrice de confusion (figure 2-4). La matrice de confusion est utilisée pour développer un ensemble de mesures de performance pour évaluer la performance des classificateurs. Toutes les mesures conventionnelles, telles que Vrai Positif (TP), Vrai Négatif (TN), Faux Positif (FP) et Faux Négatif (FN), sont déterminées à l'aide de la matrice de confusion. Les lettres TP, TN, FP et FN désignent respectivement des enregistrements qui sont identifiés avec succès comme positifs, correctement détectés comme négatifs, incorrectement détectés comme positifs et incorrectement détectés comme données négatifs.

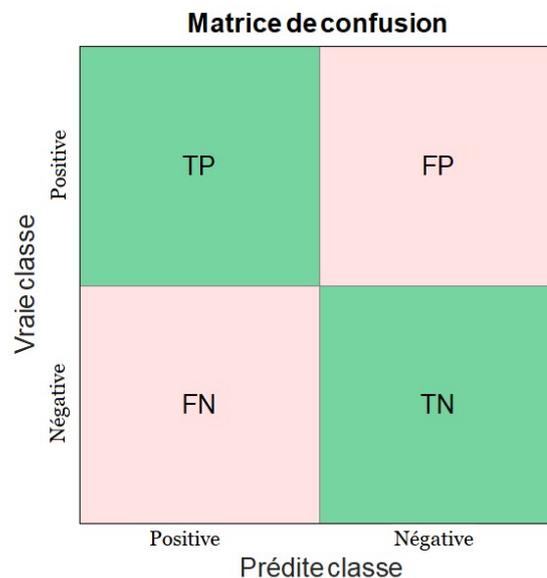


Figure 2-4 Matrice de confusion et les mesures conventionnelles

A partir de la matrice de confusion, nous pouvons tirer plusieurs indicateurs dont le plus répandu est la précision (Accuracy), c'est le rapport des classifications correctes sur le nombre total d'échantillons et il est calculé comme suit :

$$\text{Précision} = \frac{TP + TN}{TP + TN + FP + FN}$$

III.3. Méthodes de régression des données

III.3.1. Régression par moindres carrés partielles (PLSR)

Développée à l'origine pour l'économétrie dans les années soixante-dix du siècle dernier par Herman Wold [20] mais elle est devenue populaire d'abord dans la chimiométrie (c'est-à-

dire la chimie computationnelle) en particulier grâce au fils d'Herman, Svante d'après Geladi et Kowalski [21] et dans l'évaluation sensorielle [22]. La régression des moindres carrés partiels (PLSR, en anglais : Partial least Squares Regression) est une méthode supervisée, statistique et multivariée qui trouve un modèle de régression linéaire en projetant des variables prédites et observables dans un nouvel espace. Le principe du PLSR est similaire à celui de l'analyse en composantes principales (ACP) sauf que le choix des nouveaux axes ne se fait pas en fonction de la variance maximale mais en fonction des covariances maximales entre les données.

PLS est utilisé pour trouver les relations fondamentales entre deux matrices (X et Y), c'est-à-dire une approche par variable latente pour modéliser les structures de covariance dans ces deux espaces. Un modèle PLS tentera de trouver la direction multidimensionnelle dans l'espace X qui explique la direction de variance multidimensionnelle maximale dans l'espace Y [23]. Étant donné une paire de matrices de données d'entrée et de sortie X et Y et en supposant qu'elles sont liées linéairement par :

$$Y = XC + V$$

Où V et C sont respectivement des matrices de bruit et de coefficients, la régression PLS construit un modèle linéaire en décomposant les matrices X et Y en termes bilinéaires. Le modèle général sous-jacent de PLS multivarié est :

$$X = TP^T + E$$

$$Y = UQ^T + F$$

Où X est une matrice n×m de prédicteurs, Y est une matrice n×p de réponses ; T et U sont des matrices n×1 qui sont, respectivement, des projections de X (le score X, composante ou matrice factorielle) et des projections de Y (les scores Y) ; P et Q sont respectivement m×1 et p×1 matrices de chargement orthogonales ; et les matrices E et F sont les termes d'erreur, supposés être des variables normales aléatoires indépendantes et identiquement distribuées. Les décompositions de X et Y sont faites de manière à maximiser la covariance entre T et U [24].

La matrice X fait référence aux variables prédictives et à leurs carrés ou/et termes croisés s'ils ont été ajoutés. Les variables x actives participant au modèle sont souvent appelées termes. Lorsque des carrés et des termes croisés sont ajoutés à la matrice X, cela correspond à l'ajustement de modèles quadratiques. Le modèle PLS consiste en une projection simultanée des espaces X et Y sur un hyperplan de faible dimension. Les coordonnées des points sur cet

hyperplan constituent les éléments de la matrice T. Cette analyse a les deux objectifs suivants : (1) bien approximer les espaces X et Y ; (2) pour maximiser la corrélation entre X et Y.

III.3.2. Régression linéaire des Séparateurs à Vaste Marge (ϵ -SVM)

La régression des machines à vecteurs de support (ϵ -SVM) est une extension de l'algorithme SVM qui est utilisé pour prédire les valeurs [25]. Dans la régression ϵ -SVM, l'ensemble de données d'apprentissage contient des variables prédictives et des valeurs de réponse observées. Le but est de trouver une fonction $f(x)$ qui s'écarte de y_n d'une valeur ne dépassant pas pour chaque point d'apprentissage x la valeur de ϵ , et en même temps aussi plate que possible.

Supposons que nous ayons un ensemble de données d'apprentissage où x_n est un ensemble multivarié de N observations avec des valeurs de réponse observées y_n . Pour trouver la fonction linéaire ($f(x) = A\beta + C$) et confirmer qu'elle est aussi plate que possible, il faut trouver B avec la valeur de norme minimale ($\beta'\beta$). Ceci est formulé comme un problème d'optimisation convexe à minimiser :

$$J(\beta) = \frac{1}{2} * \beta' \beta$$

Sous réserve de tous les résidus ayant une valeur inférieure à ϵ .

Dans le cas de l'absence d'une fonction B qui puisse satisfaire ces contraintes pour tous les points. Ce problème peut être résolu en introduisant les variables d'écart ξ_n et ξ_n^* pour chaque point :

$$J(\beta) = \frac{1}{2} * \beta' \beta + C \sum_{n=1}^n (\xi_n + \xi_n^*)$$

Sujet à :

$$\forall n : y_n - (x_n' \beta + b) \leq \epsilon + \xi_n, (x_n' \beta + b) - y_n \leq \epsilon + \xi_n^*, \xi_n^* \geq 0, \xi_n \geq 0$$

La constante C est la contrainte de boîte, une valeur numérique positive qui contrôle la pénalité imposée aux observations situées en dehors de la marge epsilon (ϵ) et aide à éviter le surapprentissage (régularisation) [26]. Cette valeur détermine le compromis entre la planéité de B et la quantité jusqu'à laquelle des écarts supérieurs à ϵ sont tolérés.

III.3.3. Réseau Autorégressif non linéaire à entrées exogènes (NARX)

Avant d'expliquer le réseau autorégressif non linéaire à entrées exogènes (NARX, en anglais : Nonlinear Autoregressive Network with eXogenous inputs), un aperçu sur les réseaux de neurones doit être donné pour mieux comprendre le principe de fonctionnement de cet algorithme.

III.3.3.1. Généralités sur les réseaux de neurones artificielles (RNA)

En raison des différents types de problèmes qui peuvent être traités, l'utilisation des réseaux de neurones a augmenté de façon exponentielle ces dernières années. Ils peuvent être utilisés efficacement pour les problèmes de classification, de modélisation et de prédiction. McCulloch et Pitts [27] sont généralement reconnus comme les concepteurs du premier modèle mathématique d'un réseau de neurones en 1943. Ils ont créé le premier neurone formel qui calcule la somme pondérée des signaux reçus en entrée suivi d'une fonction de transfert qui répond par 0 ou 1 suivant l'état du résultat est ce qu'il est supérieur ou inférieur à un seuil [28]. La figure 2-5 montre comment le neurone formel fonctionne :

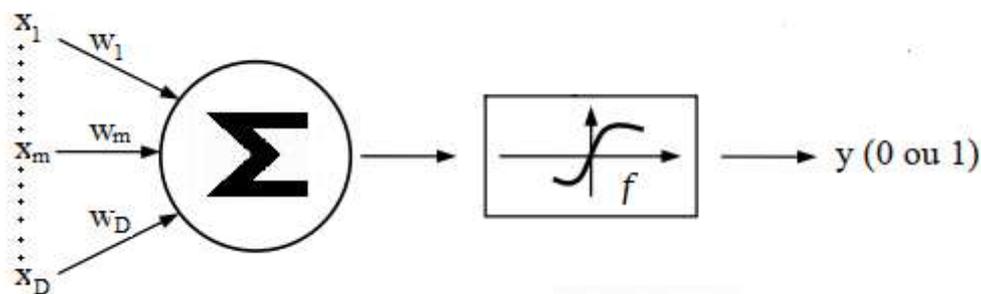


Figure 2-5 Conception d'un neurone formel d'après McCulloch et Pitts

D'après la figure, l'équation de calcul de la sortie du neurone formel est :

$$y = f \left(\sum_1^D x_i w_i - \mu \right)$$

Avec f la fonction de transfert, les x_i sont les entrées, les w_i les coefficients de pondération et μ le seuil de la fonction f .

Un réseau de neurones est composé de nombreux neurones qui sont reliés par des liaisons synaptiques qui ont un poids ou coefficient qui leur est attribué. Ces connexions agrèent

à tout neurone du réseau d'acquérir et d'envoyer des signaux à d'autres neurones. Les neurones sont organisés sous forme de couches.

Des chercheurs comme Hebb (Hebb learning) [29], Rosenblatt (Perceptrons) [30], Kohonen (Kohonen self-organizing networks) [31] et d'autres ont continué à étudier les réseaux de neurones au cours des 30 années suivant la création de ce modèle par McCulloch et Pitts, construisant des réseaux de neurones plus sophistiqués et plus variés.

Les RNA offrent un certain nombre d'avantages, notamment la capacité de gérer des ensembles de données non linéaires et une tolérance plus élevée au bruit et à la dérive des capteurs. Les RNA produisent également des résultats rapides, sont efficaces dans le traitement des données et s'entraînent par eux-mêmes. Cependant, les RNA présentent un certain nombre d'inconvénients, notamment la difficulté à déterminer les meilleurs paramètres de réseau et les meilleures techniques de formation.

Etant des méthodes non linéaires, les réseaux de neurones (RNA) sont parmi les méthodes les plus adaptées pour le traitement des données des matrices de capteurs de gaz qui génèrent des données de même nature (non linéaire). Les réseaux de neurones artificiels (RNA) sont appliqués dans les systèmes multi-capteurs ou nez électroniques depuis les années 1990 [32]. Le perceptron multicouche (PMC) est la structure la plus utilisée dans ce domaine.

III.3.3.2. Perceptron Multicouche (PMC)

Inventé par Frank Rosenblatt en 1958 [30], c'est l'un des réseaux de neurones les plus souvent utilisés pour les applications de classification et de prédiction. Il contient normalement deux ou trois couches de neurones complètement liés. Le principe fondamental du perceptron multicouche est d'organiser les neurones en couches et toutes les deux couches voisines sont totalement connectées. Par conséquent, les sorties de la première couche sont les entrées des neurones de la deuxième couche. Le vecteur d'entrée est le même pour tous les neurones de la première couche, qui sont connectés au monde extérieur. Ils calculent ensuite leurs sorties, qui sont envoyées aux neurones de la deuxième couche, et ainsi de suite. La sortie du réseau est formée par les sorties des neurones de la dernière couche. Cette couche a le même nombre de neurones que le nombre de classes prédéfinies selon le cas étudié. La figure 2-6 représente un exemple d'un réseau perceptron multicouche

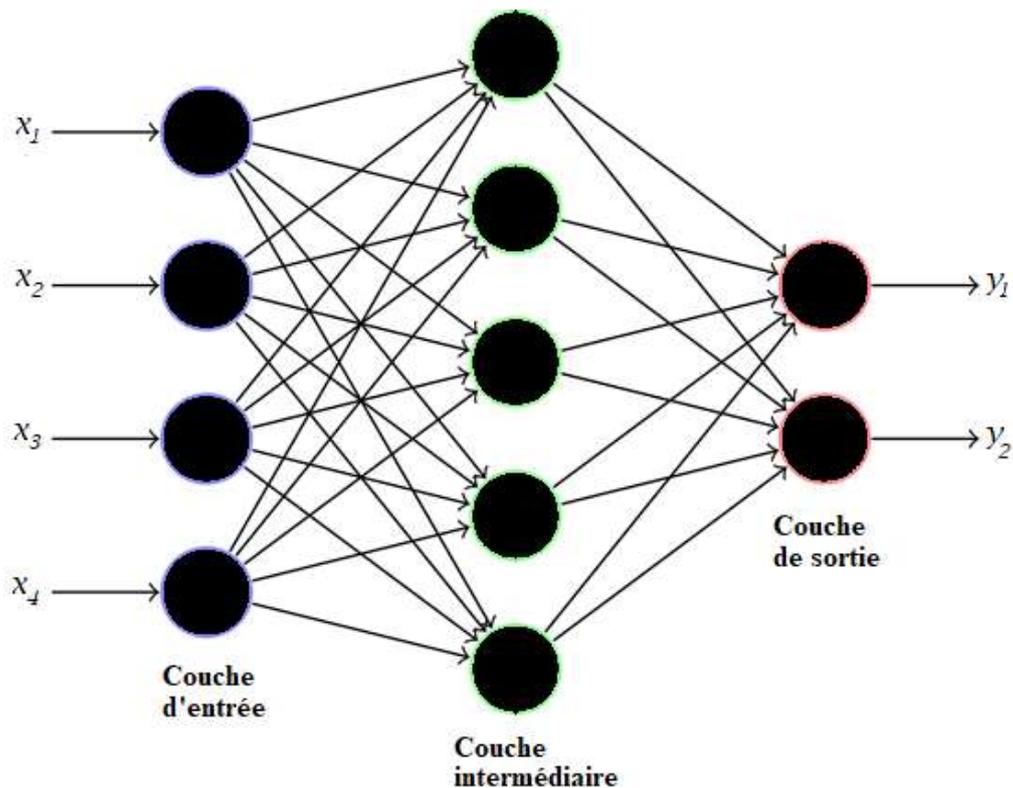


Figure 2-6 Exemple d'un réseau multicouche

Deux conditions doivent être remplies avant que les réseaux multicouches puissent être utilisés dans l'apprentissage :

- Une technique spécifique pour la sélection de la conception du réseau, telle que le nombre de couches cachées et de neurones par couche. Le nombre de neurones de la couche cachée est ajusté en fonction de la difficulté de la tâche de classification à résoudre.
- Une fois l'architecture est décidée, un algorithme détermine les coefficients et seuils synaptiques de la phase d'apprentissage à partir de l'échantillon de travail, dans le but de trouver les coefficients synaptiques qui minimisent l'erreur entre les variables d'entrée et de sortie en utilisant l'approche des moindres carrés. Ce niveau d'apprentissage est atteint en faisant des approximations répétées entre les sorties et les entrées. La rétropropagation est l'algorithme le plus largement utilisé (BP pour rétropropagation). Les valeurs des coefficients synaptiques sont initialement attribuées au hasard dans cette approche. De la sortie à l'entrée, l'erreur d'apprentissage (la différence entre les variables de sortie et d'entrée) est propagée.

Les valeurs des coefficients synaptiques sont modifiées à chaque passage pour réduire l'écart entre les variables d'entrée et de sortie.

III.3.3.3. Réseau NARX

Le réseau autorégressif non linéaire à entrées exogènes (NARX) est un réseau de neurones dynamique récurrent dont la rétroaction vient uniquement du neurone de sortie ; c'est un réseau puissant qui converge beaucoup plus rapidement et se généralise mieux que les autres réseaux de neurones récurrents conventionnels [33]. La figure 2-7 montre un exemple de réseau NARX avec deux couches cachées.

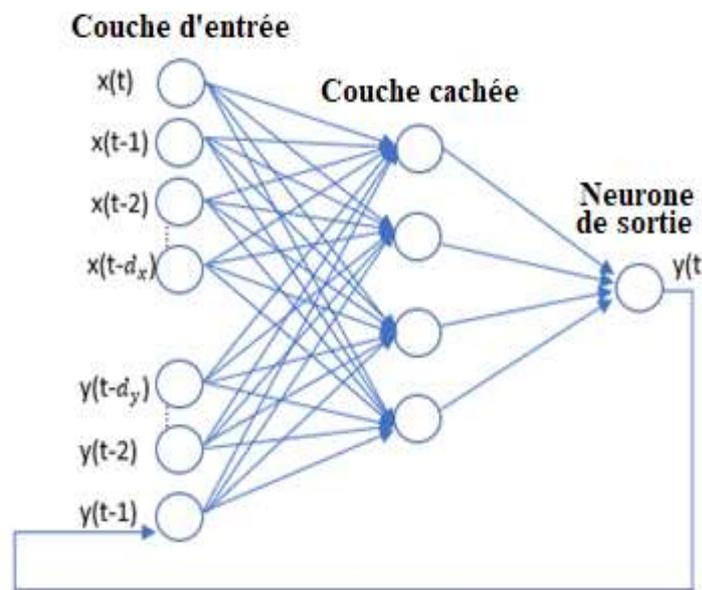


Figure 2-7 Réseau de NARX à deux couches

La figure 2-8 montre en détail la structure du réseau NARX, le premier composant est une ligne à retard à prises (TDL, en anglais Time DeLay) [34] où la série temporelle d'entrée entre par la gauche et traverse m retards. Le TDL produit un vecteur m -dimensionnel qui contient la série chronologique d'entrée au moment présent ainsi que m observations passées. La couche cachée, qui contient de nombreux neurones, est le bloc suivant. La couche de sortie, qui n'a qu'un seul neurone, est à la fin à droite. Pendant la phase d'entraînement, les poids et les biais associés aux neurones sont appris.

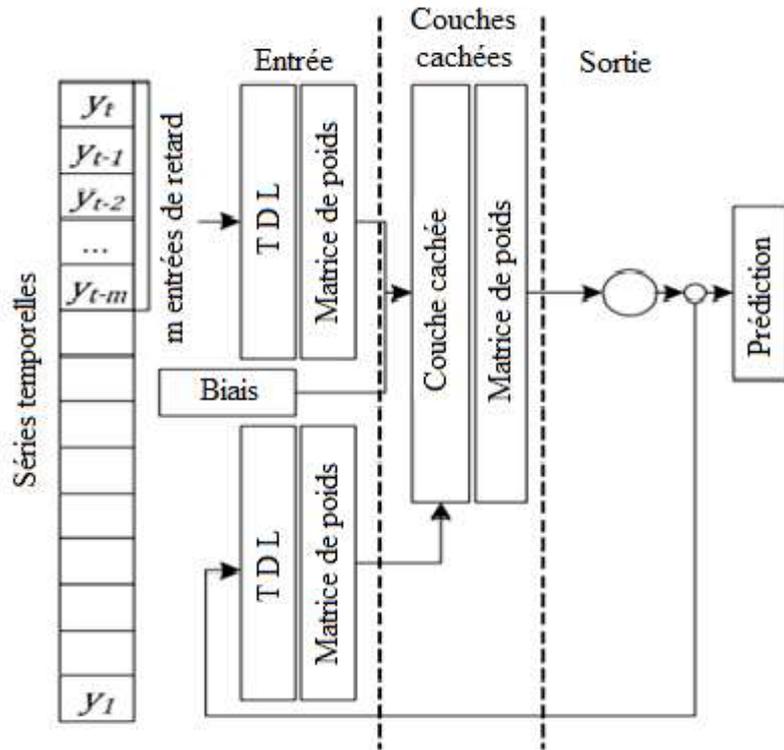


Figure 2-8 Structure du réseau NARX

Le réseau NARX peut être mis en œuvre dans l'une des deux architectures suivantes : architecture série-parallèle et architecture parallèle (figure 2-9).

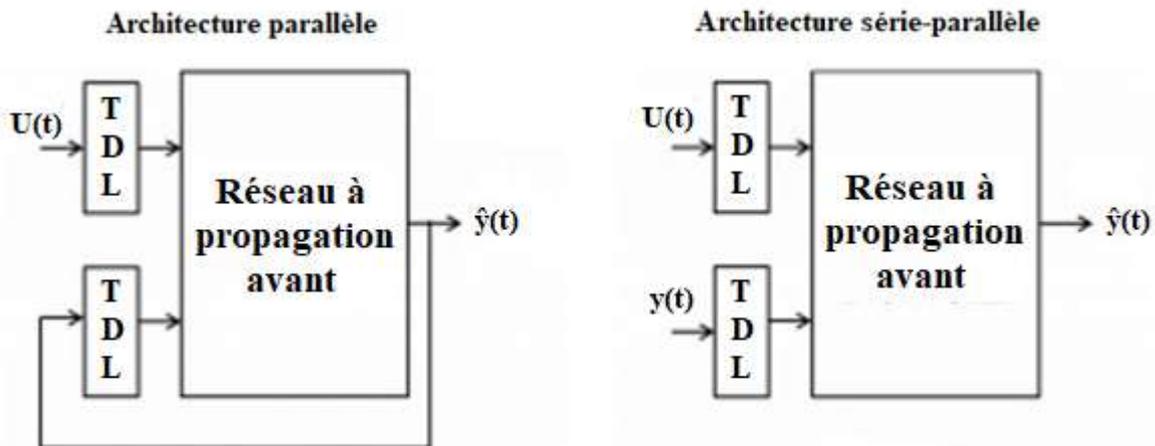


Figure 2-9 Architectures possibles du réseau NARX, avec $U(t)$ valeur de l'entrée, $y(t)$ valeur de sortie, $\hat{y}(t)$ la valeur de sortie estimée et TDL est la ligne à retard à prises

Dans l'architecture série-parallèle, la valeur future de la série temporelle $y(t - 1)$ est prédite à partir des valeurs présentes et précédentes de $x(t)$ ainsi que les vraies valeurs passées de la série temporelle $y(t)$. Dans l'architecture parallèle, la prédiction est faite en utilisant les

valeurs présentes et précédentes de $x(t)$ ainsi que des valeurs prédites passées de la série temporelle $\hat{y}(t)$.

Parce que les vraies valeurs précédentes de la série temporelle sont disponibles, l'architecture série-parallèle est utilisée dans cette étude de recherche tout au long de la phase de formation. L'utilisation de l'architecture série-parallèle présente deux avantages. Le premier avantage réside dans le fait que le réseau d'anticipation est plus précis lorsque les valeurs réelles sont utilisées comme entrée. Le second avantage consiste en une architecture du réseau résultant purement anticipative.

Après la phase de formation, le réseau de neurones NARX est converti en architecture parallèle, ce qui est bénéfique pour la prédiction multi-étape [35], [36].

L'expression mathématique du réseau NARX est définie selon l'équation suivante [37]:

$$y(t) = f(x(t-1), \dots, x(t-d_x), y(t-1), \dots, y(t-d_y))$$

Où f est une fonction non linéaire décrivant le comportement du système, $x(t)$ et $y(t)$ représentent l'entrée et la sortie du réseau à l'instant t , et d_x et d_y les ordres de mémoire d'entrée et de mémoire de sortie (retard à prises).

Pour résoudre le problème de surentraînement, le réseau utilise différents algorithmes de rétropropagation comme technique de régularisation, ses algorithmes consistent à mettre à jour les poids et les biais. La procédure d'apprentissage du réseau pourrait être basée sur la rétropropagation avec l'algorithme de Levenberg-Marquardt, le gradient conjugué à l'échelle ou l'algorithme de régularisation bayésienne.

L'algorithme de régularisation bayésienne minimise les erreurs moyennes quadratiques (MSE, en anglais : Mean Square Erreur) en utilisant une sélection précise de poids et de biais [38] selon l'optimisation de Levenberg-Marquardt. L'objectif de l'apprentissage bayésien est de déduire la distribution de probabilité a posteriori des poids compte tenu des données observées $P(w|y, \chi)$. Cela se fait en utilisant les informations contenues dans les données pour mettre à jour toute connaissance sur la valeur du poids avant d'obtenir les données, en utilisant le théorème de Bayes :

$$P(w|y, \chi) = \frac{p(y|w, \chi)p(w)}{p(y|\chi) = \int p(y|w, \chi)p(w)dw}$$

où $p(w)$ représente la distribution des poids à priori et $p(y | \chi)$ est la probabilité de y étant donné χ tandis que $p(y | w, \chi)$ est la fonction de vraisemblance, qui décrit toute information sur w contenue dans les données [39]. Cet algorithme nécessite généralement plus de temps, mais peut entraîner une bonne généralisation pour les ensembles de données difficiles ou bruyants.

III.3.4. Evaluation des performances des modèles de régression

L'évaluation des performances du modèle peut être effectuée à l'aide de deux paramètres, l'erreur quadratique moyenne (MSE) et le coefficient de corrélation (R). Cependant, dans la plupart des travaux scientifiques, les paramètres de L'erreur quadratique moyenne ($RMSE = \sqrt{MSE}$) et du coefficient de détermination ($R^2 = (R)^2$) sont utilisés. L'erreur quadratique moyenne MSE est la manière qui permet d'évaluer la performance des modèles en évaluant la différence entre les valeurs mesurées et prédites. Un MSE le plus bas fournit les meilleures performances d'approximation.

$$MSE = 1/n \sum_{i=1}^n (X_i - Y_i)^2$$

Où X est la valeur observée et Y est la valeur prédite tandis que n est le numéro de la valeur.

Pour l'analyse du meilleur ajustement linéaire entre les valeurs observées et prédites, le coefficient de corrélation R est utilisé. Ce coefficient est défini comme suit :

$$R = \frac{\sum_{i=1}^n X_i Y_i}{\sqrt{\sum_{i=1}^n X_i^2 \sum_{i=1}^n Y_i^2}}$$

C'est un indicateur qui permet de juger de la qualité de la régression, ce coefficient varie entre 0 et 1, c'est-à-dire entre une prédiction faible ($R=0$) et une prédiction forte ($R=1$) [40].

III.4. Techniques de validation des modèles

La sélection d'un algorithme de classification parmi les nombreuses méthodes est une procédure cruciale. Malgré le fait que les performances d'un classificateur peuvent être mesurées de différentes manières mais la mesure la plus courante reste la précision, qui est définie comme le pourcentage d'exemples correctement prédits divisé par le nombre total de prédictions. Cependant, lorsque l'on travaille avec un petit nombre de données, cette valeur unique n'est pas suffisamment informative car l'estimation sera totalement optimiste. Pour une

évaluation robuste de cette précision afin d'éviter le surajustement, plusieurs approches sont disponibles, mais les plus importantes sont la K-Fold validation croisée et la prédiction d'un ensemble de données indépendant appelé Hold-out Validation.

III.4.1. Simple validation

Simple validation (en anglais : Hold-out Validation) est une technique courante pour évaluer les performances des classificateurs. C'est la procédure de validation croisée la plus simple car elle implique une seule division de l'ensemble de données en deux parties. Les données de la première matrice d'échantillons (généralement les deux tiers) appelée l'ensemble d'apprentissage sont utilisées pour construire le modèle tandis que les données de la seconde matrice d'échantillons (le tiers restant) appelée l'ensemble de test sont utilisées pour évaluer la qualité de prédiction du modèle. Cependant, cette approche dépend fortement du choix des observations appartenant soit à l'ensemble d'apprentissage, soit à l'ensemble de test. Elle peut conduire à des performances sous-optimales du classificateur. Autrement dit, l'évaluation du modèle avec cette approche peut être très différente selon le découpage des données. Sans oublier qu'il y a un compromis entre la taille de la formation et l'ensemble de validation. Si une petite partie de l'ensemble de données est utilisée pour les tests, alors l'estimation de la performance de généralisation peut ne pas être fiable. D'un autre côté, lorsqu'une quantité considérable de données est utilisée pour la validation, l'apprentissage du classificateur peut être sous-optimal. Et quelle que soit la taille de la formation et des ensembles de données, il n'y a aucun moyen de savoir si les données mises de côté pour les tests ; quelle que soit sa taille, est représentatif des données que le classificateur peut voir sur les performances sur le terrain.

III.4.2. Validation croisée K-Fold

La validation croisée K-fold (en anglais : k-fold Cross Validation) fournit une estimation de la fiabilité d'un classificateur. Dans la validation K-fold, l'ensemble de données complet de N éléments est partitionné en K sous-ensembles. Parmi les K sous-ensembles, l'un est retenu comme ensemble de test et les données restantes ($K-1$) sont utilisées pour l'apprentissage du classificateur. Le processus de validation croisée est ensuite répété K fois, chacun des K sous-ensembles étant utilisé exactement une fois comme ensemble de test.

L'avantage de cette méthode est que toutes les observations sont utilisées à la fois pour l'apprentissage et la validation, et chaque observation est utilisée pour la validation une fois.

Le choix de K dépend des données. Pour des ensembles de données suffisamment volumineux, K est généralement pris comme 5 ou 10 ; cependant, pour des ensembles de données plus petits, K peut être choisi plus grand pour permettre une plus grande taille d'ensemble d'apprentissage.

III.4.3. Validation croisée Leave-one-out

Cette méthode de validation (appelée en anglais Leave-one-out cross-validation) est un cas particulier de la validation croisée K -fold avec K est égal au nombre des observations N . C'est-à-dire qu'un ensemble d'apprentissage comprenant $N-1$ observations est utilisé pour prévoir chaque membre de l'ensemble de données une fois, et N modèles sont formés pour déterminer l'erreur moyenne. Cette approche a l'avantage d'être déterministe car aucun échantillonnage aléatoire n'est utilisé, mais elle entraîne des dépenses importantes en termes de calcul.

IV. Conclusion

Dans ce chapitre, nous avons commencé par la présentation des techniques de la préparation des bases de données comme la sélection des caractéristiques et la normalisation. La sélection de caractéristiques présente un grand intérêt dans les applications du système multi-capteurs (nez électronique). En effet, elle permet d'améliorer le taux de succès de classification, de réduire la taille de la base d'apprentissage, de réduire le temps de calcul. Elle permet aussi de développer des modèles de classification simple et robuste, insensible à la dérive des capteurs et à la perturbation qui peut être dû aux fluctuations des conditions de mesure. Ensuite, nous avons présenté un nombre important de méthodes de reconnaissance de forme, pour la classification nous avons présenté ACP, SVM, KNN, arbre de décision, bayes naïf et les méthodes ensemblistes. Pour la régression, nous avons décrit PLSR, SVMR et NARX.

Références

- [1] D. F. Specht, "Rotated general regression neural network," *IEEE Trans. NEURAL NETWORKS*, vol. 2, no. 6, pp. 568–576, 1991, doi: 10.1109/IJCNN.2007.4371258.
- [2] P. C. Jurs, G. A. Bakken, and H. E. McClelland, "Computational methods for the analysis of chemical sensor array data from volatile analytes," *Chem. Rev.*, vol. 100, no. 7, pp. 2649–2678, 2000, doi: 10.1021/cr9800964.
- [3] K. Pearson, "LIII. On lines and planes of closest fit to systems of points in space,"

- London, Edinburgh, Dublin Philos. Mag. J. Sci.*, vol. 2, no. 11, pp. 559–572, 1901, doi: 10.1080/14786440109462720.
- [4] H. Hotelling, “Analysis of a complex of statistical variables into Principal Components. Jour. Educ. Psych., 24, 417-441, 498-520,” *J. Educ. Psychol.*, vol. 24, pp. 417–441, 1933.
- [5] M. Kermit and O. Tomic, “Independent component analysis applied on gas sensor array measurement data,” *IEEE Sens. J.*, vol. 3, no. 2, pp. 218–228, 2003, doi: 10.1109/JSEN.2002.807488.
- [6] A. Levin and A. Shashua, “Principal component analysis over continuous subspaces and intersection of half-spaces,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2352, pp. 635–650, 2002, doi: 10.1007/3-540-47977-5_42.
- [7] E. N. Bui, “Machine learning in the Australian critical zone,” *Data Sci. Appl. to Sustain. Anal.*, 2021, doi: Bui, E. N. (2021). Machine learning in the Australian critical zone. 10.1016/b978-0-12-817976-5.00004-8.
- [8] V. Vapnik, “Support-Vector Networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995, doi: <https://doi.org/10.1007/BF00994018>.
- [9] V. Franc and V. Hlaváč, “Vector machine multi-class support,” *Proc. - Int. Conf. Pattern Recognit.*, vol. 2, no. October 2014, pp. 236–239, 2002, doi: 10.1109/ICPR.2002.1048282.
- [10] M. A. Aizerman, E. A. Braverman, and L. Rozonoer, “Theoretical foundations of the potential function method in pattern recognition learning.,” *Automation and Remote Control*, no. 25. pp. 821–837, 1964.
- [11] C. Chang and C. Lin, “LIBSVM: A Library for Support Vector Machines,” vol. 2, no. 3, 2011, doi: 10.1145/1961189.1961199.
- [12] M. Pal, “Multiclass Approaches for Support Vector Machine Based Land Cover Classification,” 2008, [Online]. Available: <http://arxiv.org/abs/0802.2411>.
- [13] R. Debnath, N. Takahide, and H. Takahashi, “A decision based one-against-one method for multi-class support vector machine,” *Pattern Anal. Appl.*, vol. 7, no. 2, pp. 164–175, 2004, doi: 10.1007/s10044-004-0213-6.
- [14] H. Laanaya, A. Martin, D. Aboutajdine, and A. Khenchaf, “Classification des sédiments marins par fusion de classifieurs binaires SVM,” no. 1, 2006.
- [15] Z. Wang, W. Xu, J. Hu, and J. Guo, “A multiclass SVM method via probabilistic error-correcting output codes,” *Int. Conf. Internet Technol. Appl. ITAP 2010 - Proc.*, pp. 2–5, 2010, doi: 10.1109/ITAPP.2010.5566126.
- [16] B. W. Silverman and M. C. Jones, “E. Fix and J.L. Hodges (1951): An Important Contribution to Nonparametric Discriminant Analysis and Density Estimation: Commentary on Fix and Hodges (1951),” *Int. Stat. Rev. / Rev. Int. Stat.*, vol. 57, no. 3, p. 233, 1989, doi: 10.2307/1403796.

- [17] T. M. Cover, and P. E. Hart, "Nearest Neighbor Pattern Classification," vol. I, pp. 1–28, 2012.
- [18] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Second Edi. 1990.
- [19] V. B. S. Prasath, A. A. A. Haneen, A. B. A. Hassanat, O. Lasassmeh, A. S. Tarawneh, M. B. Alhasanat, H. S. Eyal Salman., "Distance and Similarity Measures Effect on the Performance of K-Nearest Neighbor Classifier -- A Review," pp. 1–39, 2017, doi: 10.1089/big.2018.0175.
- [20] V. Esposito Vinzi, Wynne W. Chin, J. Henseler, and H. Wang, *Handbook of Partial Least Squares*, vol. 206, no. 1. 2010.
- [21] P. Geladi and Bruce R. Kowalski, "PARTIAL LEAST-SQUARES REGRESSION: A TUTORIAL," *Analytica Chim. Acta*, vol. 185, pp. 1–17, 1986, doi: 10.1016/0003-2670(86)80028-9.
- [22] S. Wold and M. Sjostrom, "PLS-regression : a basic tool of chemometrics," pp. 109–130, 2001.
- [23] A. Krishnan, L. J. Williams, A. R. McIntosh, and H. Abdi, "Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review," *Neuroimage*, vol. 56, no. 2, pp. 455–475, 2011, doi: 10.1016/j.neuroimage.2010.07.034.
- [24] A. Sharifi, "Partial Least Squares-regression (PLS-regression) In Chemometrics Partial Least Squares-regression (PLS-regression) In Chemometrics," no. June 2016, 2018.
- [25] T. M. Bafithile and Z. Li, "Applicability of ϵ -Support Vector Machine and artificial neural network for flood forecasting in humid, semi-humid and semi-arid basins in China," *Water (Switzerland)*, vol. 11, no. 1, 2019, doi: 10.3390/w11010085.
- [26] VAPNIK and V. N., "The Nature of Statistical Learning," *Theory*. p. 334, 1995, [Online]. Available: <https://ci.nii.ac.jp/naid/10020951890>.
- [27] W. S. McCulloch and W. Pitts, "A logical calculus nervous activity," *Bull. Math. Biol.*, vol. 52, no. 1, pp. 99–115, 1990.
- [28] K. Ngo, "Etude d'un multicapteur pour la détection sélective des gaz," 2006.
- [29] R. G. M. Morris, "D.O. Hebb: The Organization of Behavior, Wiley: New York; 1949," *Brain Res. Bull.*, vol. 50, no. 5–6, p. 437, 1999, doi: 10.1016/S0361-9230(99)00182-3.
- [30] F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, 1958, doi: 10.1037/h0042519.
- [31] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biol. Cybern.*, vol. 43, no. 1, pp. 59–69, 1982, doi: 10.1007/BF00337288.
- [32] J. E. Haugen and K. Kvaal, "Electronic nose and artificial neural network," *Meat Sci.*, vol. 49, no. SUPPL. 1, 1998, doi: 10.1016/S0309-1740(98)90054-7.

- [33] T. Lin, B. G. Horne, P. Tiiio, C. L. Giles, and S. Member, "Learning Long-Term Dependencies in," vol. I, no. 6, pp. 1329–1338, 1996.
- [34] D. Howard and B. Mark, "Neural Network Toolbox Documentation," *Neural Netw. Tool*, p. 846, 2004.
- [35] J. Buitrago and S. Asfour, "Short-term forecasting of electric loads using nonlinear autoregressive artificial neural networks with exogenous vector inputs," *Energies*, vol. 10, no. 1, pp. 1–24, 2017, doi: 10.3390/en10010040.
- [36] A. A. Ferreira, T. B. Ludermir, and R. R. B. De Aquino, "Comparing recurrent networks for time-series forecasting," *Proc. Int. Jt. Conf. Neural Networks*, pp. 10–15, 2012, doi: 10.1109/IJCNN.2012.6252459.
- [37] H. T. Siegelmann, B. G. Horne, and C. L. Giles, "Computational capabilities of recurrent NARX neural networks," *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*, vol. 27, no. 2, pp. 208–215, 1997, doi: 10.1109/3477.558801.
- [38] F. Burden and D. Winkler, *Artificial Neural Networks: Methods and Applications*. 2008.
- [39] G. B. Kingston, M. F. Lambert, and H. R. Maier, "Bayesian training of artificial neural networks used for water resources modeling," *Water Resour. Res.*, vol. 41, no. 12, pp. 1–11, 2005, doi: 10.1029/2005WR004152.
- [40] A. Eghbalzadeh, M. Javan, M. Hayati, and A. Amini, "Discharge prediction of circular and rectangular side orifices using artificial neural networks," *KSCE J. Civ. Eng.*, vol. 20, no. 2, pp. 990–996, 2016, doi: 10.1007/s12205-015-0440-y.

**Chapitre 3 : CONCEPTION ET
RÉALISATIONS EXPÉRIMENTALES
ET METHODOLOGIE**

I. INTRODUCTION

Comme il est précisé dans le premier chapitre, le système multi-capteurs (nez électronique) est un outil inspiré des systèmes olfactifs des mammifères qu'il tente d'imiter. C'est une solution non coûteuse pour détecter et classer différents mélanges d'odeurs. Les systèmes multi-capteurs se composent typiquement d'une matrice de capteurs de gaz qui aperçoivent l'odeur et qui transforment les odeurs chimiques en signaux électriques.

Ce chapitre présentera un système multi-capteurs de gaz (nez électronique) conçu et réalisé au laboratoire à base de capteurs de gaz à oxyde métallique de commerce pour l'analyse et l'évaluation des odeurs de différents échantillons des produits agricoles pour différentes applications. En plus de la description de la conception et la réalisation de différentes parties de notre outil, à savoir la cellule de mesure contenant la matrice de capteurs, le banc d'alimentation stabilisée et la carte d'adaptation et de protection de la carte d'acquisition des données, une présentation du logiciel utilisé pour l'acquisition de données sera mise en œuvre. Nous définissons ensuite, le protocole expérimental utilisé pour réaliser la caractérisation des capteurs aux différents échantillons sujet de notre étude.

II. DISPOSITIF EXPERIMENTAL

La combinaison de la chromatographie en phase gazeuse et de la spectroscopie de masse (GC-MS) est la technique la plus utilisée pour l'identification des composés volatils dans les mélanges complexes de composés organiques [1], mais ces instruments souffrent de plusieurs inconvénients qui diminuent leurs importances comme le coût cher, la complexité de l'instrumentation qui nécessite une personne bien formée et de bonne technicité, en plus du temps nécessaire pour l'analyse sans oublier que la lourdeur de ces instruments pousse à faire des déplacements au laboratoire pour la réalisation des analyses. De ce fait, les systèmes multi-capteurs (nez électronique) sont une solution optimale, et comme déjà vu dans le premier chapitre, ces instruments se composent d'un dispositif (côté matériel) en plus d'une partie de traitement des données.

Le dispositif expérimental est un élément clé dans tout système de détection et d'analyse des odeurs, c'est un outil qui doit être conçu pour détecter et distinguer des odeurs complexes à l'aide d'un réseau de capteurs, ce dispositif peut être réalisé selon 3 configurations distinctes selon l'application et le besoin, une conception dite ouverte, statique ou dynamique.

La conception du système est une étape intéressante car la configuration choisie affecte, par la suite, les résultats des analyses.

II.1. Types de conceptions des systèmes nez électroniques

II.1.1. Dispositif ouvert

Simple dans sa conception, le dispositif ouvert [2] est une configuration qui consiste à placer une matrice de capteurs dans un environnement ouvert sans besoin d'échantillonnage, il repose sur la diffusion des espèces odorants vers la matrice pour la reconnaissance des échantillons (la figure 3-1). Sûrement, c'est un dispositif très facile à réaliser, cependant il présente de nombreux inconvénients car la reconnaissance des espèces gazeux, sera très difficile et peut être faussée à cause de l'exposition aux nombreuses grandeurs interférentes comme d'autres gaz ou polluants contenus dans l'air (parfums, fumées de cigarettes, poussières...) ou encore les paramètres atmosphériques (variations de l'humidité, de la température ou de la vitesse de l'effluent) qui perturbent la mission des capteurs. Ce type de configuration est utilisé principalement pour la surveillance de l'environnement.

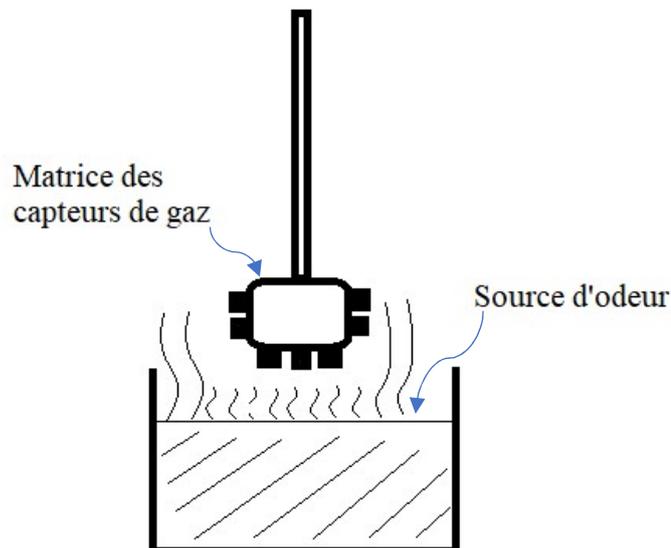


Figure 3-1 Conception ouverte d'un nez électronique

II.1.2. Dispositif statique

Le dispositif statique [3] peut être utilisé lorsque les échantillons de mélanges gazeux peuvent être obtenus et stockés avant le test. Il se compose d'une enceinte hermétique (en verre,

époxy, téflon ou d'acier inoxydable) contenant en plus de la matrice de capteurs, un agitateur et une arrivée d'échantillons par l'intermédiaire d'une seringue (la figure 3-2).

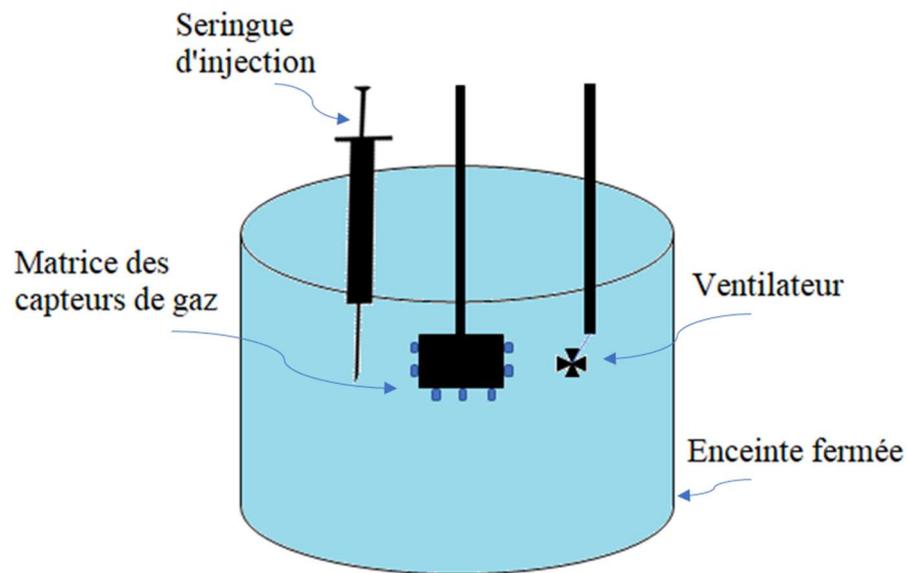


Figure 3-2 Conception statique d'un nez électronique

Pour le principe de fonctionnement de ce type de dispositif, l'enceinte contient un gaz vecteur, et à l'injection de l'échantillon gazeux, l'agitateur homogénéise le mélange qui pourra être étudié par la matrice de capteurs.

Parmi les avantages de ce type de dispositif, nous trouvons le faible coût et la simplicité de sa conception. Malgré cela, le choix de matériau de fabrication de l'enceinte reste l'une des principales restrictions. Pour éviter toutes sortes d'interactions possibles qui peuvent fausser les mesures, l'enceinte doit être construite de façon à ce que ses murettes intérieures soient lisses et ne facilitent aucun phénomène d'adsorption et de désorption de gaz, d'autant plus qu'il n'y a pas d'arrivée continue de gaz [4]. Ainsi, des flacons en verre sont fréquemment utilisés [5]. Néanmoins, des matériaux tels que l'époxy, le téflon ou l'acier inoxydable ont aussi été utilisés [6]–[8].

Un autre problème à ne pas négliger est lié à la taille de l'enceinte, car le grand volume signifie une grande consommation de temps pour chaque expérience, ceci est dû aux périodes de propagation et d'équilibre très longues [9].

Un autre problème cette fois-ci lié au mode opératoire, c'est la préparation des échantillons de test à introduire dans l'enceinte, il faut garder le même protocole de préparation

des dilutions afin de garantir la répétitivité et par conséquent la fiabilité des résultats acquis et obtenus.

La présence de ces inconvénients qui nuisent à l'utilisation de ce genre de dispositif ne limite pas son adoption, portant il est utilisé dans les analyses d'odeurs dans différents domaines incluant l'agroalimentaire [10].

II.1.3. Dispositif dynamique

La conception de ce genre de dispositif (figure 3-3) se compose d'une enceinte - contenant la matrice de capteurs de gaz - dotée de deux conduites, une pour l'entrée du flux de gaz pompé avec un ventilateur ou par un gaz vecteur, ce qui indique une exposition continue incessante aux flux gazeux pendant un intervalle de temps qui peut être prédéfini [11]. La conduite du flux de gaz peut être doté d'un débitmètre pour contrôler le débit afin de garantir les mêmes conditions de traitement sauf si un ventilateur à débit constant est utilisé.

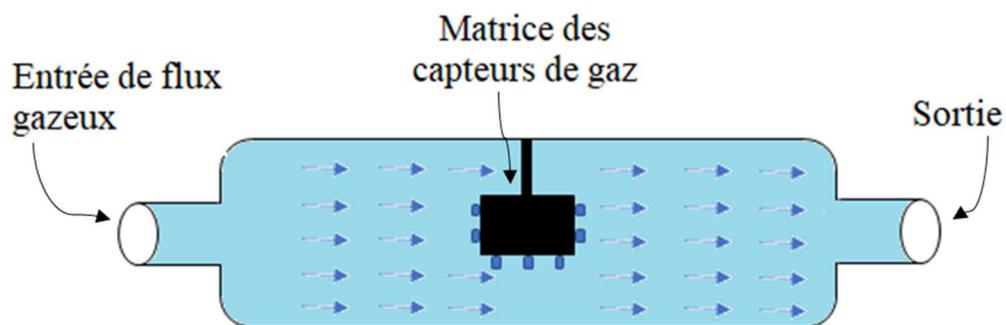


Figure 3-3 Conception dynamique d'un nez électronique

En comparant ce type de dispositif à celui présenté avant (statique), le présent offre plusieurs avantages. Certes, en traitant le mélange d'odeurs comme un flux continu, la décomposition gazeuse du mélange provoquée par des réactions chimiques avec la couche sensible des capteurs serait continuellement vidée et remplie par le mélange d'odeurs d'origine à examiner. De plus, ce type de dispositif permet de résoudre le problème des réactions d'adsorption et de désorption par les parois de l'enceinte, et se caractérise par sa rapidité et offre la possibilité de contrôler les paramètres expérimentaux et de faire des mesures crédibles et reproductibles.

Le dispositif dynamique est généralement le dispositif le plus largement adopté dans les applications de caractérisation des capteurs de gaz [12]. Dans ce travail, le dispositif dynamique est la conception que nous avons préféré pour la réalisation de notre outil de mesure pour l'étude des odeurs générées par les produits agroalimentaires vu les multiples avantages qu'il présente. Notre dispositif est constitué par quatre parties principales : un système d'échantillonnage de la substance odorante à analyser composé d'un ventilateur et de la chambre de l'échantillon, l'enceinte de mesure qui contient la matrice de capteurs sensibles aux composés odorants, un système de conditionnement de signal et d'acquisition de données, une partie logicielle permettant le prétraitement de données pour la classification et l'identification des odeurs.

II.2. Conceptions et réalisations

II.2.1. Choix des capteurs

Ce travail vise à concevoir un outil de contrôle des aliments traités fiable et moins coûteux pour détecter les pesticides et la nature de certains produits agricoles en l'occurrence la menthe et les pommes de terre.

Pour la menthe, 77 composés volatils ont été identifiés par Y. J. Park et coll. [13] dans *Mentha Spicata L.*, les plus présents sont les composés organiques volatils comme le limonène et des alcools. Cette substance vitale serait étudiée traitée avec deux insecticides parmi les plus utilisés par les agriculteurs contenant deux matières toxiques à savoir le malathion et le deltaméthrine. Une suspension aqueuse obtenue par la dilution dans l'eau de malathion commercial a été étudiée par N. Suto et H. Kawashima [14] en utilisant la chromatographie en phase gazeuse/spectrométrie de masse à rapport isotopique a montré que cette matière libère des composés organiques volatils comme l'éthylbenzène et le xylène. Tandis que F. Aiello et coll. [15] ont étudié le deltaméthrine qui est un composé ester dont la structure de base est constituée d'un alcool aromatique et d'un acide cyclopropylcarboxylique.

Pour les pommes de terre, J. F. Dresow and H. Böhm [16] ont analysé des tubercules de pommes de terre en substances volatiles qui dégagent dans les trois cas : crues bouillies et cuites au four et ils sont arrivés à identifier 159 volatils dans les pommes de terre crues dont nous trouvons des alcools, des hydrocarbures, des composés soufrés, des composés azotés et d'autres composés organiques volatils comme l'acétone, l'undécane, le benzène et le limonène. En plus, le découpe ou le tranchement de ce légume provoque une attaque oxydative et enzymatique sur les lipides inhérents ce qui conduit à une grande quantité d'aldéhydes et d'alcools [17]. Tandis

que les analyses faites sur les fumiers de poulets ont dévoilé leurs compositions qui comptent 72 composés [18] qui ont été identifiés par la chromatographie en phase gazeuse et la chromatographie en phase gazeuse-spectrométrie de masse, parmi ces volatils nous trouvons des alcools, des esters d'acide, des alcanes, des acides carboxyliques, les Alcan-amides, les composés soufrés et des aldéhydes.

Présents au marché et connus par leurs sensibilités ainsi que leur faible coût et leur fiabilité, ces critères nous ont conduits à l'utilisation des capteurs à oxydes métalliques de commerce. La différence entre les capteurs présentés par les différentes sociétés se situe au niveau de l'élément sensible même si la plupart utilisent le dioxyde d'étain SnO_2 , il existe d'autres capteurs à base des oxydes de tungstène WO_3 , de zinc ZnO , ou de titane TiO_2 . Une autre différence a été remarquée, c'est au niveau de la conception interne du capteur qui se reflète sur le boîtier ainsi nous trouvons des capteurs à 2, 4 ou encore 6 broches (figure 3-4).

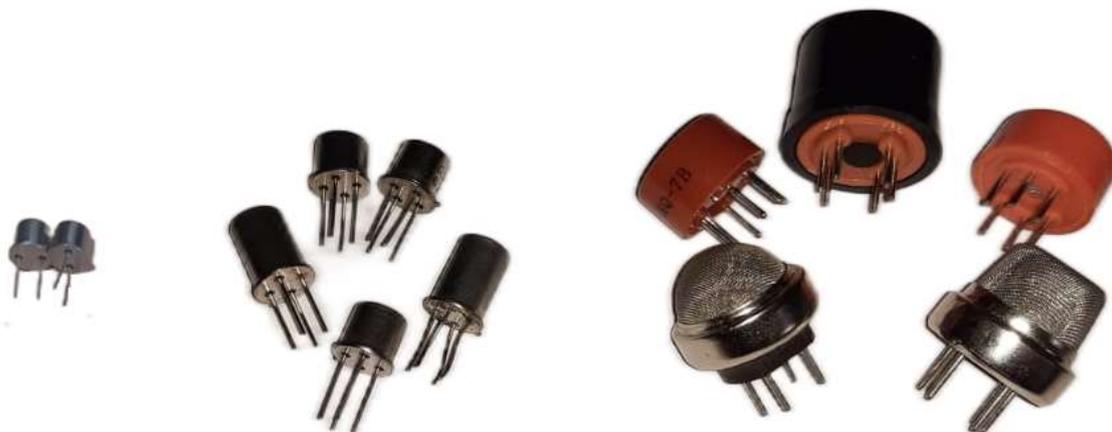


Figure 3-4 Exemple des différents capteurs présents dans les marchés

Pour chaque capteur de gaz à oxyde métallique, la mise en contact avec différents composés odorants volatils induit une variation de la conductivité du matériau sensible à base d'oxydes métalliques semi-conducteurs. Ces réactions sont différentes pour chaque capteur, donnant à chaque odeur une signature olfactive unique qui lui est propre.

Après l'étude des composés volatils des matières, sujet de notre étude, il s'est avéré qu'il existe de nombreux gaz que nous pouvons cibler, intrinsèques aux produits l'objet de l'étude et aux produits même du traitement, c'est pourquoi nous avons préféré de concevoir un support doté de sept places pour l'insertion d'une variété de capteurs que nous avons choisi en se basant

sur l'étude des composés volatils dans un premier temps pour pouvoir arriver à créer des signatures olfactives des produits étudiés. La raison était le manque de sélectivité des capteurs de gaz à base d'oxyde métallique, et la seule méthode fiable pour y remédier est de s'appuyer sur la sensibilité croisée en utilisant une matrice de plusieurs capteurs différents (différents types et/ou couches sensibles). En effet, l'emploi de capteurs d'architectures différentes dans une matrice permet d'engendrer des réponses variées pour un gaz spécifique du fait de leur cinétique d'adsorption variable.

Nos capteurs à base d'oxydes métalliques choisis sont conçus par deux fabricants très répandus dans le marché : sept de type TGS de la société japonaise Figaro et quatre de type MQ de la société chinoise Hanwei.

Les capteurs utilisés sont sensibles aux gaz odorants, aux vapeurs d'alcool, aux composés organiques volatils...etc. Le tableau 3-1 suivant cite les capteurs achetés et utilisés :

Tableau 3-1 Capteurs utilisés, leurs gaz cibles et prix

CAPTEURS	GAZ CIBLES	PRIX ESTIME (DOLLAR AMERICAIN)
TGS821	Hydrogène, méthane, éthanol et monoxyde de carbone	74
TGS822	Composés organiques volatiles (COV), éthanol, fuite de gaz, alcool	14
TGS826	Gaz odorants tels que l'ammoniac, air, isobutane, hydrogène et éthanol	80
TGS2620	Composés organiques volatiles (COV), alcool	11
TGS2611	Hydrogène, méthane, éthanol, isobutane et air	11
TGS2602	Gaz odorants tels que l'ammoniac et le sulfure d'hydrogène	9

TGS4161	Dioxyde de carbone	28
MQ-7	Monoxyde de carbone, hydrogène, méthane et air d'alcool	1.5
MQ-136	Sulfure d'hydrogène, ammoniac, air et monoxyde de carbone	14.5
MQ-137	Ammoniac, oxyde de benzène, monoxyde de carbone et air	20
MQ-214	Méthane, Air, Isobutane et propane	2

II.2.2. Matrice de capteurs

Les nez électroniques sont généralement composés des composants matériels (partie hardware) conjointement avec un traitement de données (partie software). Dans la partie hardware, la matrice des capteurs est le composant principal jouant le rôle du récepteur qui sera en contact direct avec le mélange gazeux, et grâce à ses réactions, les décisions vont être prises. C'est pour cela que la sélection des capteurs et la bonne conception est très intéressante.

La matrice de capteurs est hébergée à l'envers dans une cellule de mesure fabriquée en verre avec une couverture en plastique entouré par du caoutchouc qui évite la sortie de l'air et qui assure une bonne distribution de pression et débit de gaz, cette cellule de petite taille aide à diriger les échantillons gazeux vers les capteurs. La matrice de capteurs conçue est réalisée en circuit imprimé sur une carte électronique PCB implanté sur la couverture de la cellule (Figure 3-5). La conception du circuit a été faite à l'aide du logiciel TCI (voir annexe).

La chambre des capteurs (la cellule) est dotée d'une entrée de gaz qui est percée au-dessus sur la couverture en traversant aussi la carte des capteurs. L'entrée de cette dernière est liée à l'alimentation de commande V_c et de chauffage V_h par un connecteur mâle d'une nappe de 20 connexions ainsi qu'à la sortie où elle est connectée à une carte d'adaptation et de protection de la carte d'acquisition de données.

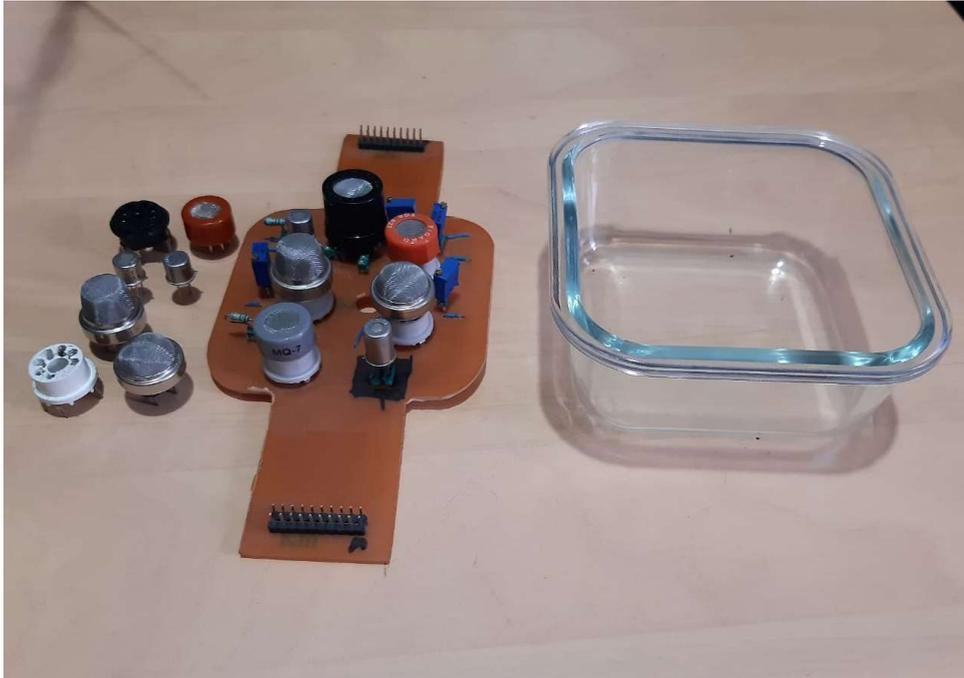


Figure 3-5 Matrice de capteurs réalisée

II.2.3. Alimentation de la matrice des capteurs

Le plus souvent les capteurs de gaz à oxyde métalliques nécessitent deux sources de tension continues, une pour la mesure de résistance de la couche sensible, et l'autre pour alimenter la résistance de chauffage à l'exception des capteurs à deux pattes comme le capteurs MQ-214. La couche sensible des capteurs doit être chauffée à des valeurs de 150 °C et plus pour répondre aux gaz cibles et le non-respect des tensions de chauffage va influencer les réponses. Les montages électroniques correspondant aux types de capteurs choisis sont présentés par la figure 3-6.

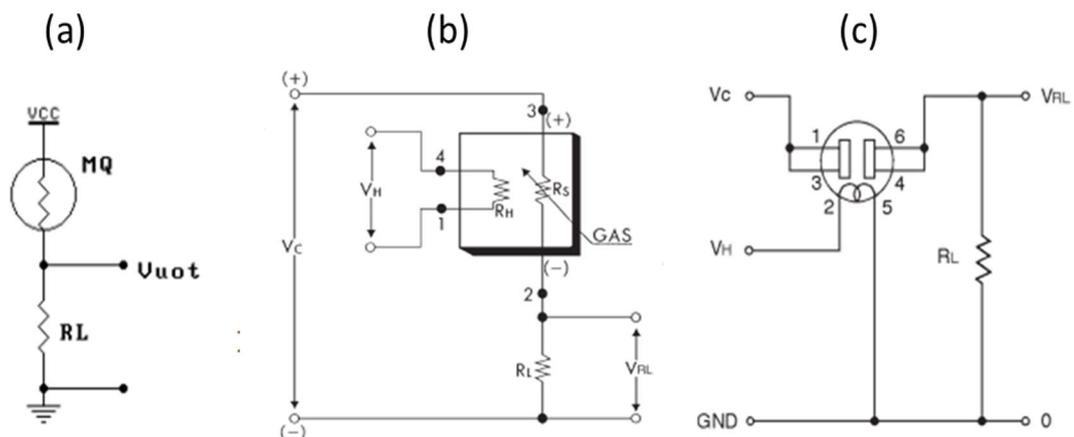


Figure 3-6 Montages électroniques correspondant aux types de capteurs choisis, (a) pour les capteurs à 2 broches, (b) pour les capteurs à 4 broches et (c) pour les capteurs à 6 broches

Dans cette figure on trouve :

- Pour les capteurs à deux broches :
 - V_{CC} : tension de l'alimentation du circuit,
 - V_{OUT} : tension de sortie,
 - R_L : résistance de mesure,
- Pour les capteurs à quatre broches :
 - V_C : tension de l'alimentation du circuit,
 - V_H : tension de chauffage,
 - V_{RL} : tension mesurée aux bornes de la résistance de charge R_L ,
 - R_L : résistance de mesure,
 - R_H : résistance de chauffage,
 - R_s : résistance de la couche sensible,
- Pour les capteurs à six broches :
 - V_C : tension de l'alimentation du circuit,
 - V_H : tension de chauffage,
 - V_{RL} : tension mesurée aux bornes de la résistance de charge R_L ,
 - R_L : résistance de mesure,
 - Entre les broches 2 et 5 : résistance de chauffage.

La figure 3-6 (a) représente le schéma de circuit d'un capteur à deux broches. La particularité de ce capteur réside dans l'utilisation d'une seule source de tension V_{cc} qui sert à la fois pour l'alimentation et le chauffage de l'élément sensible du capteur. En présence de l'odeur, le capteur adsorbe l'oxygène, sa résistance interne R_s change. Par conséquent, les fluctuations de la tension V_{OUT} sont détectées.

$$V_{OUT} = V_{cc} \times \frac{R_L}{R_L + R_s}$$

Avec R_s est la résistance interne du capteur

La figure 3-6 (b) et (c) illustrent successivement les schémas des circuits d'un capteur à quatre et à six broches, avec une simple différence c'est que dans le cas d'un capteur à six broches, la première broche est shuntée à la troisième et la quatrième est aussi shuntée à la sixième. La tension de source est fournie aux circuits via V_C et à l'élément chauffant via V_H . Les résistances des deux capteurs R_s changent lorsque les capteurs adsorbent l'oxygène. Ainsi

les variations des tensions V_{OUT} sont mesurées pendant que les capteurs adsorbent l'odeur testée. La concentration du gaz désoxydant affecte les résistances des capteurs. La relation entre les deux peut être exprimée par l'équation suivante :

$$R_S = \left(\frac{V_C}{V_{RL}} - 1 \right) \times R_L \quad \text{Et} \quad V_{RL} = \left(\frac{R_L}{R_L + R_S} \right) V_C$$

De ce fait, un banc de l'alimentation stabilisée conçu et réalisé est présenté dans l'annexe. Il est doté de trois sorties de 5V, une sortie variable arrivant jusqu'à 10V, une sortie de +12V et une autre de -12V pour l'alimentation des amplificateurs opérationnels, et quatre bornes des sorties pour la masse.

II.2.4. Système d'acquisition de données

II.2.4.1. Carte d'acquisition de données ADLINK USB-1901

Les cartes d'acquisition de données jouent le rôle d'interface entre l'ordinateur et les matrices des capteurs pour acquérir des résultats. Elles sont le lien entre les deux composantes du nez électronique, la partie réceptrice contenant principalement la matrice des capteurs de gaz et la dernière et importante partie software qui sert au traitement des données pour la reconnaissance de l'odeur.

À l'aide des cartes d'acquisition de données, les signaux analogiques sont transformés en sortie numérique qu'on peut sauvegarder sur un ordinateur pour un traitement antérieur.

La carte d'acquisition choisie est DAQ ADLINK USB-1901 (figure 3-7), elle appartient à la série ADLINK USB-1900 qui propose une gamme de modules DAQ multifonctions basés sur USB 2.0. C'est une carte d'acquisition reconnue par ses avantages qui se manifestent dans la simplicité, la portabilité, la fiabilité et la flexibilité. Elle permet une configuration précise, rapide et simple en fournissant un conditionnement de signal intégré, une alimentation USB et une connectivité USB plug-and-play.

La carte d'acquisition DAQ ADLINK USB-1901 permet quatre plages d'entrée de tension différentes (± 10 V, ± 2 V, ± 1 V, ± 200 mV) avec 16 entrées analogiques de résolution 16 bits avec taux d'échantillonnage maximum qui peut arriver jusqu'à 250 kS/s, 8 entrées numériques et 4 sorties numériques TTL. Les entrées de cette carte sont protégées par une haute impédance qui dépasse 1 Giga ohm. Cette carte peut tolérer une température de fonctionnement allant de 0°C jusqu'à 55°C. Elle est alimentée par un bus USB équipé de bornes à vis amovibles

pour une connectivité facile des appareils. Un support multifonctionnel est fourni avec la carte qui peut être utilisé pour l'attacher pour un montage sur un bureau, sur un rail ou sur un mur. Cette carte d'acquisition est adaptée aux tests de signaux mixtes, à la recherche en laboratoire et à l'automatisation d'usine.



Figure 3-7 Carte d'acquisition de données ADLINK USB-1901

II.2.4.2. Carte d'adaptation et de protection

Comme son nom l'indique, la carte d'adaptation et de protection a été conçue pour une double fonction : l'adaptation et la protection.

Après des tests de mesure, la carte d'acquisition de données n'a pas réussi à transmettre les signaux des capteurs pour les visualiser sur l'ordinateur. Après la vérification de la datasheet de la carte, il s'est avéré que les entrées de la carte sont protégées par une grande résistance qui dépasse $1\text{ G}\Omega$ ce qui provoque les chutes de tensions à ses bornes. La solution a été d'utiliser un amplificateur opérationnel comme amplificateur tampon pour l'adaptation d'impédance, c'est pour empêcher la source de signal d'être affectée par les courants (ou tensions, pour un tampon de courant) que la charge peut être produite avec. Et pour la miniaturisation le choix est tombé sur le TL082CP qui contient deux amplificateurs dans un boîtier de 8 broches.

La carte est aussi destinée à protéger la carte d'acquisition contre les surtensions car la tension maximale pouvant être appliquée aux entrées analogiques de la carte d'acquisition est

limitée à 10 V comme indiqué par le constructeur. C'est pour cela que nous avons conçu, pour chaque capteur, un simple circuit de protection constitué d'une diode Zener d'une tolérance de tension de 4.7V, et une résistance de 220Ω. La figure 3-9 représente le montage adopté et la figure 3-8 représente la réalisation réelle de la carte d'adaptation et de protection.

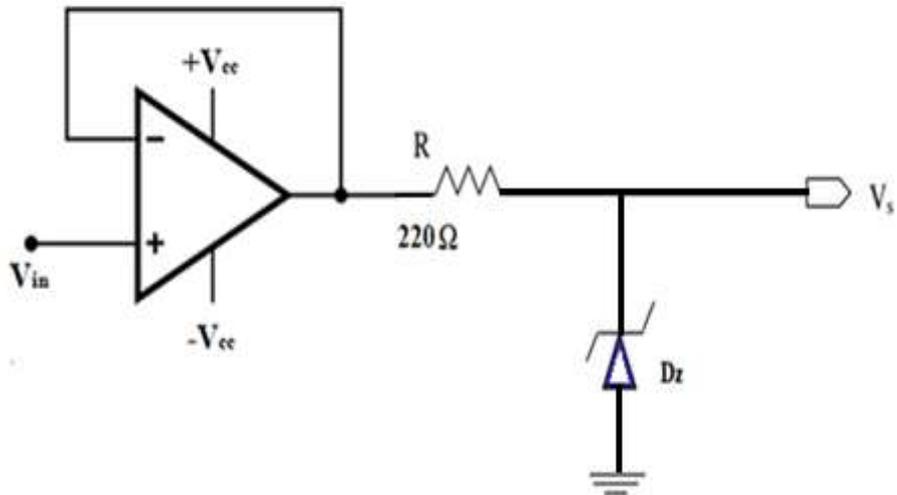


Figure 3-8 Montage d'adaptation et de protection pour chaque capteur

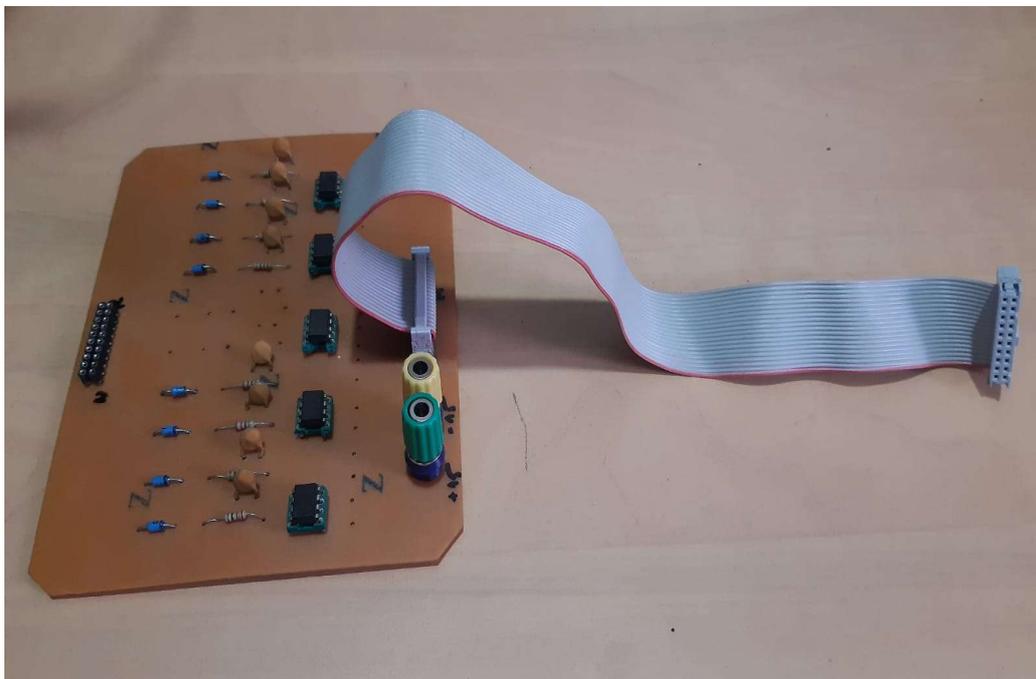


Figure 3-9 Carte d'adaptation et de protection réalisée

II.2.4.3. Description du logiciel utilisé LabVIEW

LabVIEW est l'abréviation de Laboratory Virtual Instrument Engineering Workbench, c'est un logiciel de programmation graphique « G » doté d'une vaste bibliothèque de fonctions mathématiques et statistiques utilisé pour développer des applications d'acquisition, d'analyse et de présentation des données analogiques et numériques. LabVIEW aide à créer des instruments virtuels (VIs) dans un bloc diagramme grâce à son langage de programmation graphique G.

Dans notre cas, le progiciel LabVIEW utilisé est celui de 2017, il a été exploité pour collecter les réponses de la matrice des capteurs de gaz sous forme de données analogiques. Grâce à sa face-avant les données ont été contrôlées visuellement avant d'être enregistrées à la fin de chaque expérience sous format Excel.

La figure 3-10 représente le schéma fonctionnel de la programmation graphique LabVIEW utilisé dans notre cas pour l'acquisition des données, et la figure 3-11 représente la face-avant du dit programme.

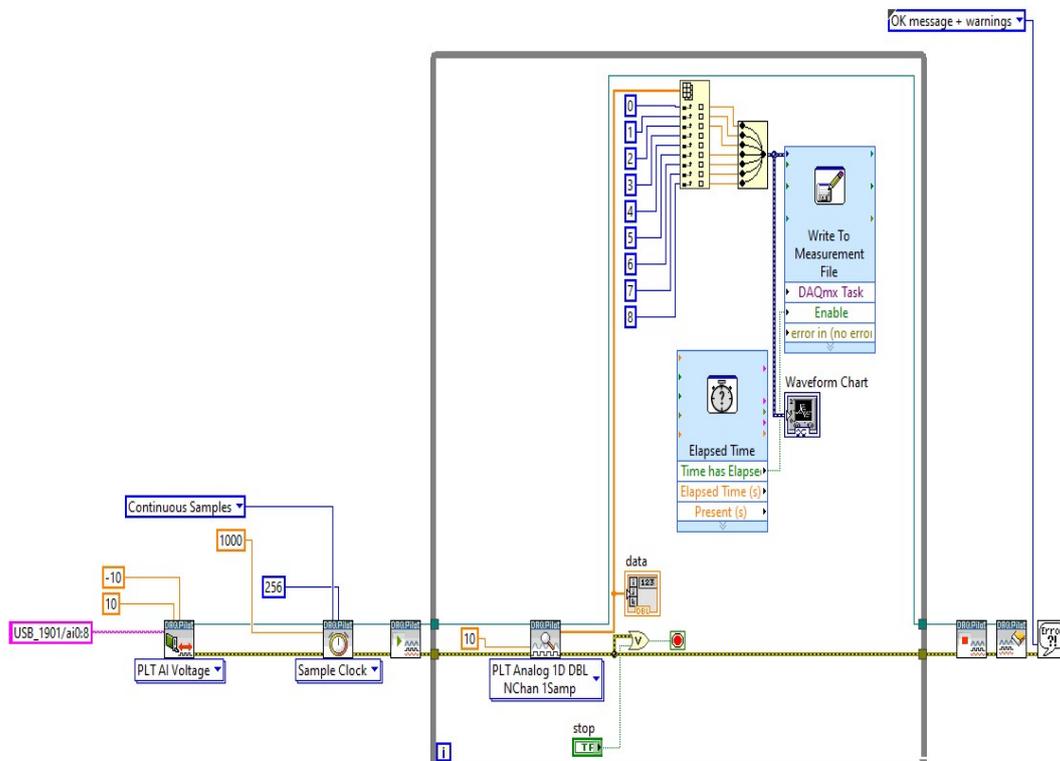


Figure 3-10 Schéma fonctionnel du programme LabVIEW réalisé et utilisé pour l'acquisition de données

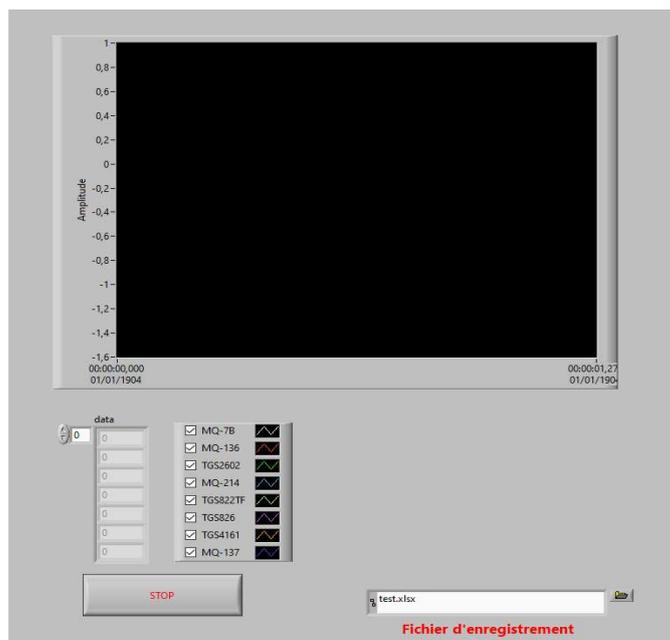


Figure 3-11 Face-avant du programme LabVIEW d'acquisition de données

II.3. Présentation de l'ensemble du dispositif expérimental

II.3.1. Schéma global

Le système de mesure réalisé dans notre laboratoire est schématisé sur la figure 3-12, il est polyvalent pour différentes études et applications.

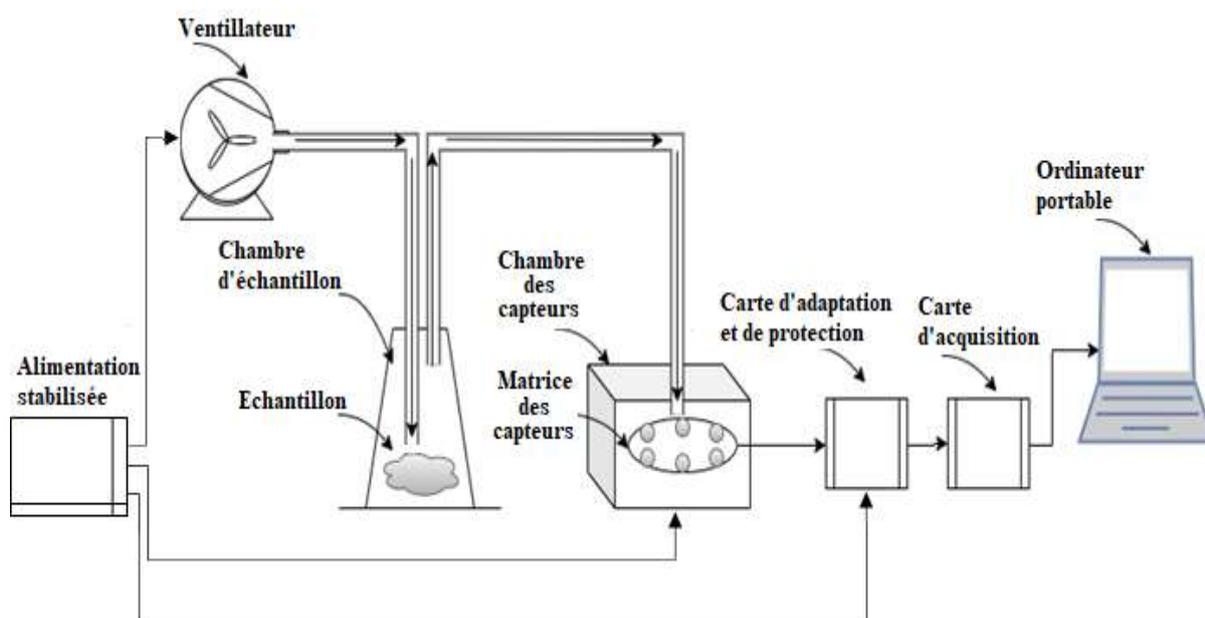


Figure 3-12 Représentation schématique de l'outil utilisé

Notre système se compose d'un ventilateur, d'une chambre d'échantillonnage, d'une chambre de capteurs qui contient la matrice de capteurs, d'un circuit de conversion et de protection, d'une carte d'acquisition de données (ADLINK USB 1901 DAQ) et d'un ordinateur personnel PC.

Le circuit du réseau de capteurs contient également des résistances de charge pour ajuster la sensibilité du capteur, en plus de deux connexions de nappe de 20 pins pour les entrées des alimentations et les sorties des réponses.

Une fois le système électronique développé est réalisé, il a été testé pour s'assurer qu'il fonctionne de manière satisfaisante. Les composants volatils propagés dans l'espace échantillon de la menthe présentent un changement dans l'une des caractéristiques électroniques des capteurs qui se manifeste par une augmentation ou une diminution de la tension, du courant ou de la résistance en fonction de la concentration du gaz.

II.3.2. Conditions expérimentales

- Tension d'alimentation et de chauffage de 5V,
- Durée de préchauffage de deux jours pour la stabilisation des capteurs,
- Flux gazeux stable généré par un ventilateur à volume de flux d'air maximal de 10 CFM,
- Durée de passage du mélange gazeux de 8 min,
- Durée de repos en air ambiant des capteurs de 10 min.

II.3.3. Procédure de mesure

Le système se compose d'une chambre avec un réseau de capteurs qui reçoit un flux d'air contenant l'odeur des échantillons provenant de la chambre d'échantillonnage poussé par un ventilateur. L'espace de tête de l'échantillon est injecté dans la chambre du capteur pendant un certain temps. Des expériences préliminaires doivent être effectuées pour vérifier le bon fonctionnement de l'entité. L'outil doit être alimenté deux jours avant le début du processus de mesure et doit toujours rester alimenté pendant les jours d'expérimentation, cela permet aux capteurs de se stabiliser et ainsi d'améliorer la reproductibilité des mesures.

Les analyses des odeurs quantifiées grâce à notre dispositif doivent être répétées un certain nombre de fois pour vérifier la répétabilité de la réponse à la même odeur.

L'acquisition des données se fera à l'aide de la carte d'acquisition DAQ ADLINK USB 1901 et du logiciel LabVIEW. Les réponses des capteurs sont enregistrées sous forme Excel avec le nom du fichier qui correspond à la nature de l'échantillon et l'ordre de l'expérience et sont stockées dans un dossier sur l'ordinateur pour un accès ultérieur lors de l'analyse des données.

II.3.4. Logiciel de traitement de données

L'étude d'un mélange de gaz à l'aide d'un système multi-capteurs ou nez électronique est certainement un problème d'analyse multivariée, étant donné que les données générées par les capteurs de la matrice sont multidimensionnelles et à grande échelle. Ces données sont fréquemment non linéaires. Après la phase de l'acquisition vient la phase du traitement de données. En conséquence, un logiciel de traitement de données multivariées performant et adéquat est nécessaire. Le logiciel Matlab version 2019a a été notre choix vu les nombreuses fonctionnalités qu'il présente :

- Il offre la possibilité de faire à la fois la recherche qualitative et quantitative ;
- Il est équipé par une variété de topologies d'algorithme, de réseaux et d'algorithmes d'apprentissage ;
- Il possède des graphiques et des outils de visualisation de données pour créer des visualisations désirées ;
- Il couvre l'ajustement des courbes, la catégorisation des données, l'analyse des signaux et une variété d'autres tâches spécialisées ;
- Il est doté de boîtes à outils (Toolboxes) supplémentaires adaptées à différents besoins individuels.

Dans nos études et grâce à ce logiciel, nous avons fait le prétraitement des données : la mise à l'échelle, l'extraction des caractéristiques et la normalisation. Et nous avons codé des algorithmes des quelques méthodes comme l'analyse en composantes principales (ACP) et les séparateurs à vaste marge (SVM). Ainsi, nous avons utilisé la boîte à outils (toolbox) Machine Learning and Deep Learning et spécifiquement les applications :

- Classification Learner ;
- Regression Learner ;
- Neural Net Time Series.

III. Conclusion

Dans ce chapitre, nous avons présenté les types des conceptions des dispositifs expérimentaux de nez électroniques : ouvert, statique et dynamique. En prenant compte des applications envisagées, nous avons justifié notre choix de la conception de type dynamique, c'est la conception la plus adaptée à notre cas qui présente de meilleurs avantages pour l'obtention d'une détection odorante reproductible et fiable. Nous avons présenté ainsi les différentes parties formant l'outil conçu et réalisé.

Enfin, nous avons donné la schématisation globale représentative du système de mesure réalisé et les conditions expérimentales qui doivent être respectées pour le bon déroulement des expériences ainsi que la procédure de mesure.

Références

- [1] A. L. Makas and M. L. Troshkov, "Field gas chromatography-mass spectrometry for fast analysis," *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.*, vol. 800, no. 1–2, pp. 55–61, 2004, doi: 10.1016/j.jchromb.2003.08.054.
- [2] P. Corcoran and H. V. Shurmer, "An intelligent gas sensor," *Sensors Actuators A. Phys.*, vol. 41, no. 1–3, pp. 192–197, 1994, doi: 10.1016/0924-4247(94)80110-X.
- [3] E. Llobet, J. Brezmes, X. Vilanova, J. E. Sueiras, and X. Correig, "Qualitative and quantitative analysis of volatile organic compounds using transient and steady-state responses of a thick-film tin oxide gas sensor array," *Sensors Actuators, B Chem.*, vol. 41, no. 1–3, pp. 13–21, 1997, doi: 10.1016/S0925-4005(97)80272-9.
- [4] F. Di Francesco, M. Falcitelli, L. Marano, and G. Pioggia, "A radially symmetric measurement chamber for electronic noses," *Sensors Actuators, B Chem.*, vol. 105, no. 2, pp. 295–303, 2005, doi: 10.1016/j.snb.2004.06.013.
- [5] J. W. Gardner, E. L. Hines, and M. Wilkinson, "Application of artificial neural networks to an electronic olfactory system," *Meas. Sci. Technol.*, vol. 1, no. 5, pp. 446–451, 1990, doi: 10.1088/0957-0233/1/5/012.
- [6] P. Rekha, M. S. Suresh, and V. Kusanur, "Sensor for Measuring Aroma of Jasmine," *Proc. IEEE Sensors*, vol. 2018-Octob, pp. 1–4, 2018, doi: 10.1109/ICSENS.2018.8589921.
- [7] Z. Wang, W. Chen, S. Gu, Y. Wang, and J. Wang, "Evaluation of trunk borer infestation duration using MOS E-nose combined with different feature extraction methods and GS-SVM," *Comput. Electron. Agric.*, vol. 170, no. February, p. 105293, 2020, doi: 10.1016/j.compag.2020.105293.
- [8] R. C. Young, W. J. Buttner, B. R. Linnell, and R. Ramesham, "Electronic nose for space program applications," *Sensors Actuators, B Chem.*, vol. 93, no. 1–3, pp. 7–16, 2003, doi: 10.1016/S0925-4005(03)00338-1.

- [9] P. Mielle and F. Marquis, "Alternative way to improve the sensitivity of electronic olfactometers," *Sensors Actuators, B Chem.*, vol. 58, no. 1–3, pp. 526–535, 1999, doi: 10.1016/S0925-4005(99)00158-6.
- [10] J. Brezmes, B. Ferreras, E. Llobet, X. Vilanova, and X. Correig, "Neural network based electronic nose for the classification of aromatic species," *Anal. Chim. Acta*, vol. 348, no. 1–3, pp. 503–509, 1997, doi: 10.1016/S0003-2670(97)00202-X.
- [11] H. V. Shurmer and J. W. Gardner, "Odour discrimination with an electronic nose," *Sensors Actuators B. Chem.*, vol. 8, no. 1, pp. 1–11, 1992, doi: 10.1016/0925-4005(92)85001-D.
- [12] I. Harvey, G. Coles, and J. Watson, "The development of an environmental chamber for the characterization of gas sensors," *Sensors and Actuators*, vol. 16, no. 4, pp. 393–405, 1989, doi: 10.1016/0250-6874(89)85009-7.
- [13] Y. J. Park, T. B. Baskar, S. K. Yeo, M. V. Arasu, N. A. Al-Dhabi, S. S. Lim and S. U. Park "Composition of volatile compounds and in vitro antimicrobial activity of nine *Mentha* spp.," *Springerplus*, vol. 5, no. 1, 2016, doi: 10.1186/s40064-016-3283-1.
- [14] N. Suto and H. Kawashima, "Classification of nine malathion emulsion samples by using carbon isotope ratios and the ratio of organic solvents," *Sci. Justice*, vol. 57, no. 1, pp. 1–5, 2017, doi: 10.1016/j.scijus.2016.10.009.
- [15] F. Aiello, M. G. Simons, J. W. van Velde, and P. Dani, "New insights into the degradation path of deltamethrin," *Molecules*, vol. 26, no. 13, 2021, doi: 10.3390/molecules26133811.
- [16] J. F. Dresow and H. Böhm, "The influence of volatile compounds of the flavour of raw, boiled and baked potatoes: Impact of agricultural measures on the volatile components," *Landbauforsch. Volkenrode*, vol. 59, no. 4, pp. 309–338, 2009.
- [17] M. A. Petersen, L. Poll, and L. M. Larsen, "Changes in Flavor-affecting Aroma Compounds during Potato Storage Are Not Associated with Lipoxygenase Activity," *Am. J. Potato Res.*, vol. 80, no. 6, pp. 397–402, 2003, doi: 10.1007/BF02854251.
- [18] A. Yasuhara, "Identification of volatile compounds in poultry manure by gas chromatography-mass spectrometry," *J. Chromatogr. A*, vol. 387, no. C, pp. 371–378, 1987, doi: 10.1016/S0021-9673(01)94539-X.

**Chapitre 4 : RESULTATS ET
DISCUSSIONS : ETUDE DE CAS DE
LA MENTHE COMMESTIBLE**

Introduction

Ce premier chapitre des résultats et discussions est consacré à l'étude de la menthe comestible, il présentera des expériences réalisées en utilisant le nez électronique conçu et réalisé au sein du laboratoire afin de discriminer la menthe non traitée de la menthe traitée avec des insecticides.

La menthe est une plante aromatique de la famille des Lamiacées qui prospère dans les zones humides. Elle compte 18 espèces différentes mais trois d'entre elles avec leurs variantes prédominent le marché : *Mentha Aquatica* (menthe aquatique), *Mentha Canadensis* (menthe du Canada), et *Mentha spicata* (menthe verte) [1].

Au Maroc, la consommation de la menthe (spécialement *Mentha spicata*) est très fréquente puisque le thé à la menthe est une boisson nationale traditionnelle, il est également utilisé comme plante médicinale et dans la préparation de produits cosmétiques grâce aux vertus de ses huiles essentielles, il occupe une superficie d'environ 3 500 hectares et fournit en moyenne 50 000 tonnes [2]. La culture de la menthe est pratiquée dans toutes les régions du Maroc, mais certaines villes sont reconnues par la grande production et la qualité de la récolte en l'occurrence Tiznit, Meknès, et El Brouj.

La menthe est également victime de problèmes phytosanitaires dû aux ravageurs, aux maladies cryptogamiques et virales [3], causant ainsi de grandes pertes aux agriculteurs et réduisant leur marge bénéficiaire. Et pour protéger leurs cultures, les agriculteurs utilisent certains types d'insecticides qui doivent normalement disparaître au bout de quelques jours. D'autre part, la demande des consommateurs augmente ce qui pousse les agriculteurs à ne pas respecter le délai de pré-récolte. C'est pour cette raison que certains aliments de consommation contiennent des résidus d'insecticides.

Dans une première partie nous allons utiliser notre système nez électronique conçu afin de discriminer entre les échantillons de menthe prises à partir de trois champs différents : un champ non traité, un autre traité avec l'insecticide Decis Fluxx et un autre traité avec l'insecticide Malyphos 50, cette préliminaire étude a été réalisée sur les quatre premiers jours après la pulvérisation du dits produits.

Pour la deuxième partie, nous se focaliserons, cette fois, sur la discrimination de la menthe non traitée de celle traitée avec le Malyphos 50 ainsi que la prédiction du jours de traitement de la menthe traitée en étalant notre étude sur sept jours.

I. Première partie : DETECTION DES INSECTICIDES DANS LA MENTHE COMESTIBLE

L'idée de cette étude a été lancée après avoir appris que l'Office National Marocain de Sécurité Sanitaire des Aliments (ONSSA) avait découvert que les résultats d'analyses - faites au laboratoire – des échantillons de la menthe, présentaient des taux de non-conformités très élevés, dû soit à l'utilisation de produits chimiques non agréés, soit à des taux élevés de résidus des produits agréés.

La surveillance de la qualité des aliments est essentielle compte tenu de l'utilisation intensive d'insecticides qui menacent la santé humaine [4]. Pour la détection des insecticides dans les échantillons alimentaires et agricoles comme : le malathion et la deltaméthrine qui sont l'objet de nos études, nous trouvons l'utilisation des méthodes onéreux et compliquées. La plupart des méthodes s'articulent sur l'utilisation de la chromatographie [5] : la chromatographie en phase gazeuse couplée à la spectrométrie de masse (CG-SM) [6], [7], la chromatographie en phase liquide couplée à la spectrométrie de masse en tandem (CL-SM/SM) [8], la chromatographie en phase gazeuse avec détection par capture d'électrons (CG-DCE) [9] ou avec détection par ionisation de flamme (CG-DIF) [10]. Soit la chromatographie liquide à haute performance couplée à l'ultraviolet (CLHP-UV) [11], [12] ou à la spectrométrie de masse en tandem (CLHP -SM/SM) [13], [14].

Cependant, ces méthodes sont ennuyeuses et nécessitent des instruments coûteux ainsi qu'un temps de réponse long. Le présent travail a pour objectif de faciliter la tâche de surveillance de qualité de la menthe (traitée ou non traitée) en proposant un outil simple, rapide et non destructif.

I.1. Préparation des échantillons

La première étape de notre étude était de cultiver de la menthe naturelle, à cet effet nous avons planté de la menthe (*Mentha spicata L.*) dans un terrain libre de 9 mètres carrés dans notre établissement, l'Ecole Nationale des Sciences Appliquées de Houribga, Maroc. L'objectif était de s'assurer que nous travaillons avec de la menthe non traitée. Après que la menthe ait bien mûri, nous avons divisé notre petit champ en trois parties distinctes comme suit

: la première est maintenue naturellement sans traitement, la seconde pulvérisée avec l'insecticide Decis Fluxx et la troisième pulvérisée avec l'insecticide Malyphos 50. Le choix des insecticides a été fait sur la base de travaux de recherche [2], [15] menés dans différentes régions du Maroc où l'enquête a montré l'utilisation de deux produits de la plupart des cultivateurs de menthe, le Decis Fluxx contenant de la deltaméthrine [16] qui est un composé chimique de la famille des pyréthroïdes neurotoxiques utilisés contre les insectes ou les serpents, sa formule chimique est $C_{22}H_{19}Br_2NO_3$. Le Malyphos 50 contient du malathion [17] qui est un composé organophosphoré cancérigène neurotoxique interdit dans plusieurs pays sa formule chimique est $C_{10}H_{19}O_6PS_2$.

Une émulsion en spray a été préparée juste avant la pulvérisation au dosage préconisé par le vendeur, soit 2mL dans un litre d'eau pour le Decis Fluxx et 4mL dans un litre d'eau pour le Malyphos. Pour la pulvérisation, nous avons utilisé un litre du mélange pour pulvériser chaque partie juste une seule fois.

I.2. Réponses temporelles des capteurs

Après des tests préliminaires dans lesquels nous avons utilisé les différents capteurs acquis, il s'est avéré que seuls les capteurs TGS-822, TGS-2620 et MQ-136 ont bien réagi à la présence de différents types d'échantillons de la menthe. Après ces tests, nous avons passé à la réalisation des expériences, 15 expériences par jour divisées sur les trois types de la menthe (5 expériences par type) durant 4 jours ont été faites comme suit : les échantillons avec un poids de $5g \pm 0,1g$ ont été prélevés frais et placés dans la chambre du capteur, à partir de laquelle ils ont subi un flux d'air émis par le ventilateur pendant huit minutes avec un temps de 30 secondes avant de déclencher le flux d'air pour observer la stabilité des capteurs. Les réactions des capteurs ont été enregistrées à l'aide de la carte d'acquisition ADLINK USB 1901 DAQ et le logiciel LabVIEW, chaque expérience a pris une vingtaine de minutes au total.

La figure 4-1 illustre les réponses en tension relative V_r en fonction du temps des trois capteurs en présence des trois types d'échantillons de la menthe (non traitée, traitée avec Decis Fluxx et traitée avec Malyphos 50) pendant quatre jours à partir du premier jour de l'application des insecticides.

$$V_r = \frac{V_m - V_0}{V_0}$$

Où V_r est la tension relative résultante, V_m est la tension mesurée et V_0 est la tension au point de départ.

Les figures (a), (b), (c) et (d) représentent les réponses en tension relative en fonction du temps en minute du capteur TGS-822 à partir du premier jour d'applications des insecticides au quatrième jour, les figures (a'), (b'), (c') et (d') sont celles du capteur TGS-2620 et les figures (a''), (b''), (c'') et (d'') sont celles du capteur MQ-136.

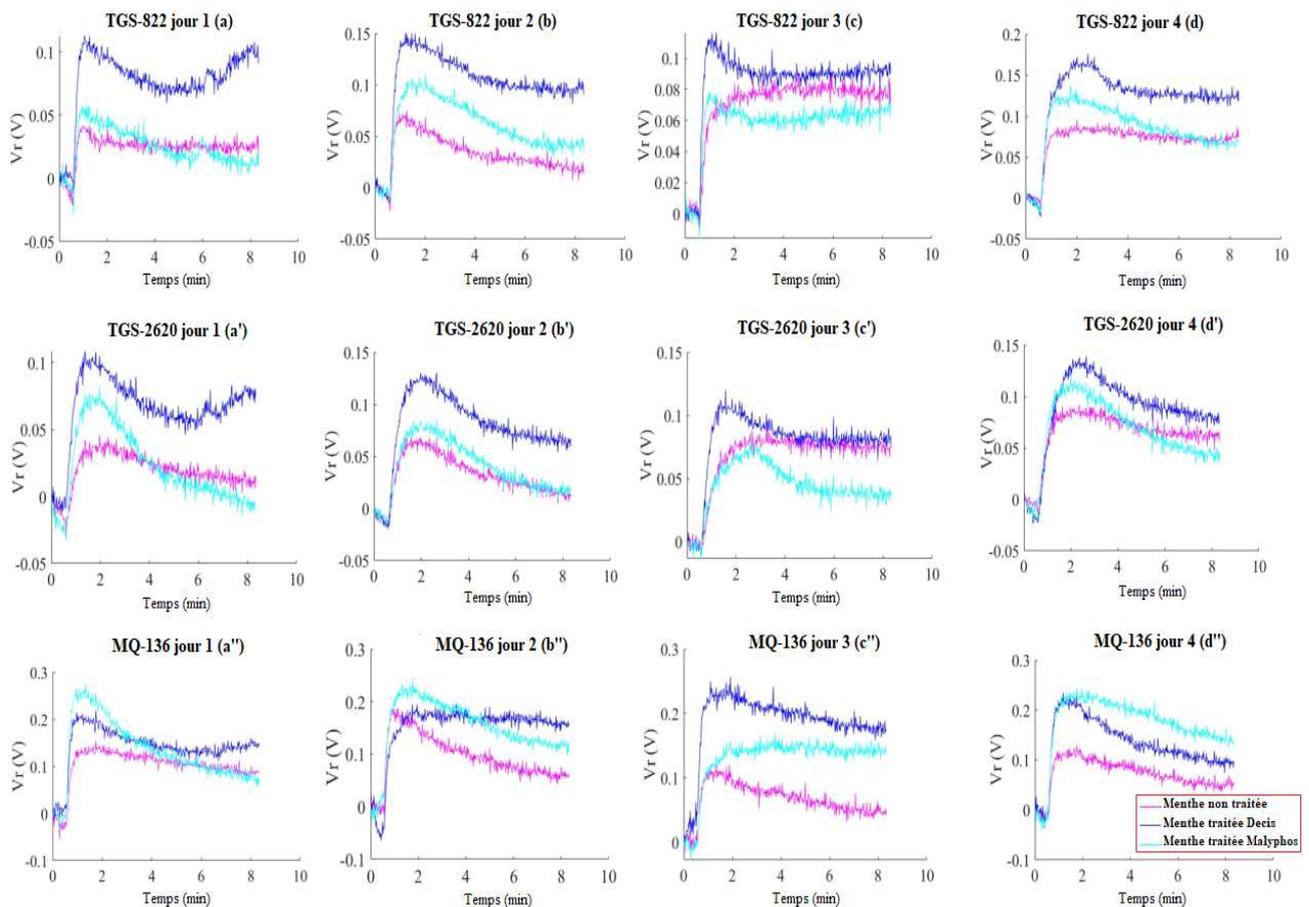


Figure 4-1 Réponse des capteurs en présence de différents types de menthe

Les composants volatils propagés dans l'espace échantillon de la menthe présentent un comportement global d'augmentation de la tension détectée à la sortie des capteurs avec une différence remarquable entre les réponses des trois différents échantillons (non traités, traités avec Decis et ceux traités avec Malyphos).

En observant les signaux de réponse des capteurs TGS-822 et TGS-2620 (figure 4.1. a, b, c, d, a', b', c' et d') on constate qu'ils ont une réponse similaire. Les signaux relatifs à la

menthe traitée avec le produit Decis Fluxx ont une valeur maximale de tension plus élevée que celles de la menthe traitée avec Malyphos 50 et celles de la menthe non traitée pendant les quatre jours, de plus, ces signaux se stabilisent à un niveau plus élevé que les autres. Aussi, les valeurs maximales des signaux de réponse relatifs à la menthe traitée au Malyphos 50 sont également relativement supérieures à celles de la menthe non traitée, par contre, les niveaux de stabilisation de ces dernières sont supérieurs à ceux de la menthe traitée au Malyphos 50.

En examinant les graphes (a''), (b''), (c'') et (d'') de la figure. 4.1, nous remarquons que les valeurs maximales des signaux relatifs à la menthe traitée par le Malyphos 50 sont les plus importantes (sauf dans le troisième jour). De plus, les valeurs maximales les plus faibles et les niveaux de stabilisation des signaux les plus faibles sont ceux de la menthe non traitée. Également, les niveaux de stabilisation du signal les plus élevés au cours des trois premiers jours est celui de la menthe traitée avec Decis Fluxx.

La différence qui existe entre les réponses des capteurs est certainement due à la différence de nature des substances volatiles libérées par chaque type de menthe, beaucoup d'informations ont été observées et remarquées dans les signaux de réponse des capteurs par leur sensibilité croisée. Il peut être utilisé pour faire la distinction entre la menthe non traitée et la menthe traitée en utilisant les algorithmes d'analyse et de traitement des données.

I.3. Discrimination de la menthe traitée de celle non traitée

Avant de passer aux analyses multivariables des données, il est à signaler que dans **tous les résultats** présents dans ce mémoire nous avons adopté un prétraitement comme suit :

- Un centrage de donnée pour supprimer les décalages par la formule :

$$x = \frac{x_m - x_0}{x_0}$$

x_m est la valeur mesurée et x_0 est la valeur au point de départ.

- La mise à l'échelle n'a pas été faite puisque toutes les variables de l'ensemble de données sont mesurées par la même unité.
- Les caractéristiques extraites sont les suivantes :
 - V_{max} : la tension maximale (valeur maximale du signal).
 - M_s : la moyenne stabilisée du signal dans l'intervalle de temps allant de 450s à 500s.
 - Aire : la zone occupée par le signal de 40s à 240s.

- La normalisation a été faite par colonne en divisant les données par la valeur maximale

$$x_{ij} = \frac{x_{ij}}{\max(x_j)}$$

Après l'étape du prétraitement des données, les données résultantes ont été organisées en quatre ensembles de données :

- X1 : Relatif à la menthe non traitée
- X2 : Relatif aux menthes traitées par Decis fluxx.
- X3 : Relatif à la menthe traitée au Malyphos 50.
- X4 : Ensemble total de données contenant toutes les données.

Chacun des jeux données X1, X2 et X3 est composé de neuf colonnes (3 réponses des capteurs × 3 caractéristiques principales) et vingt lignes (5 échantillons (considérés comme des individus) × 4 nombres de jours). Tandis que X4 est composé de neuf colonnes (3 réponses des capteurs × 3 caractéristiques principales) et soixante lignes (5 individus × 3 types de menthe × 4 nombres de jours).

I.3.1. L'analyse en composantes principales (ACP) à deux dimensions

L'analyse en composantes principales est une méthode largement appliquée avec les nez électroniques pour identifier les gaz dans tous les domaines [18] y compris l'agriculture [19]. Dans notre étude, elle a été appliquée sur les ensembles X1 et X2, X1 et X3, X2 et X3 et enfin X4 pour réduire l'espace de données de neuf dimensions à un espace de deux (PC1 et PC2). L'objectif est de visualiser les données dans un graphe de dimensions réduites, et chercher une possible distinction entre les groupes sur laquelle nous pouvons se baser pour déterminer la capacité de notre outil à distinguer entre différents types de menthe. La Figure 4-2 représente les résultats d'identification des individus des ensembles de données sur les deux premières composantes principales PC1 et PC2.

La figure 4-2 (a) représente la projection des deux ensembles de données X1 et X2 relatifs à la menthe non traitée et traitée avec le produit Decis Fluxx. Nous observons une bonne séparation entre les scores (les valeurs des individus sur les composantes principales) de chacun de ces deux groupes ainsi qu'une bonne corrélation entre les scores de chaque groupe (issus d'un même espace de tête). Les deux composantes PC1 et PC2 représentent 85 % de la variance des données c'est-à-dire que 85% de l'information contenue dans les ensembles de données X1 et X2 a été retenue et représentée. En effet, 67.53% de la variance est représentée sur le premier

axe PC1 et c'est celui qui permet d'identifier principalement la menthe non traitée de la menthe traitée avec Decis Fluxx.

La figure 4-2 (b) illustre la projection des individus des deux ensembles de données X1 et X3 correspondant à la menthe non traitée et celle traitée avec le produit Malyphos 50, un chevauchement entre les deux groupes est présent, avec une bonne corrélation entre les scores de la menthe non traitée par rapport à ceux de la menthe traitée qui sont un peu dispersés. Le taux de contribution de la variance cumulée de PC1 et PC2 est d'environ 89 %. Même si l'axe PC2 ne contribue que par 12,74% de la variance, il contribue principalement dans la séparation entre la menthe non traitée et la menthe traitée avec Malyphos 50.

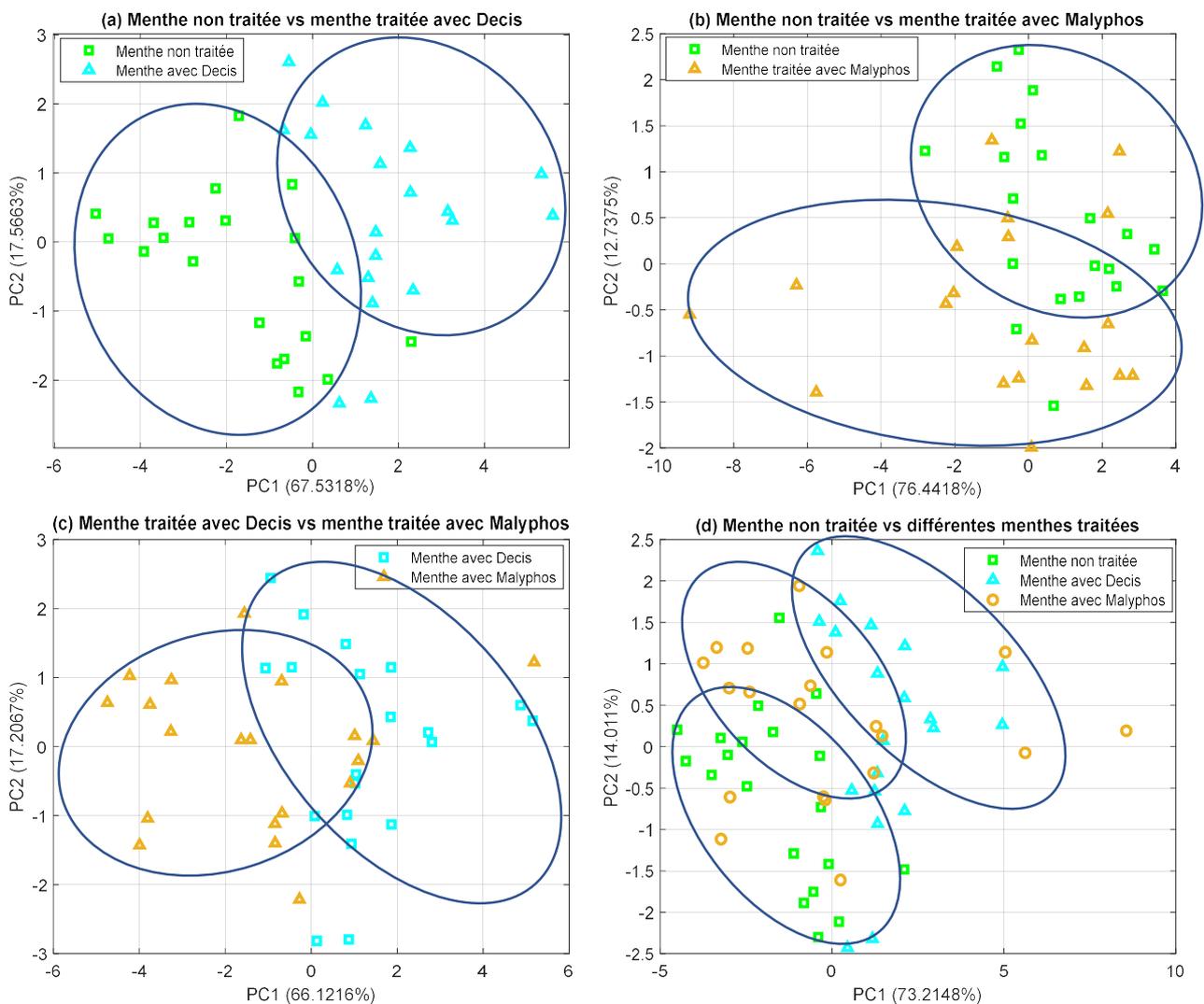


Figure 4-2 Résultats de l'ACP des données de capteurs obtenues à partir de : (a) ensemble de données X1 et X2, (b) jeux de données X1 et X3, (c) ensembles de données X2 et X3 et (d) ensembles de données globales X4

La figure 4-2 (c) représente la projection des deux ensembles de données X2 et X3 correspondant à la menthe traitée avec le produit Decis Fluxx et la menthe traitée avec le produit Malyphos 50. Nous observons que les scores des individus de la menthe traitée avec Malyphos 50 sont corrélés entre eux tandis que ceux de Decis fluxx sont dispersés. Les deux groupes se chevauchent légèrement, et le taux de contribution de variance cumulée est 83 %. Le premier axe PC1 représente 66,12% de la variation, et c'est celui qui distingue, le plus, le type du traitement qui a subi la menthe. Tandis que la figure 4-2 (d) représente les données des trois types de menthe à la fois chevauchées avec une variance de données de 87 %.

Sur la base des résultats de la méthode ACP qui offre une classification visuelle des individus des ensembles de données, les représentations des individus issus de différents types de menthe nous permettent d'estimer les similitudes et les différences entre les mesures effectuées. Par conséquent, il apparaît que la menthe traitée avec le produit Decis Fluxx se distingue facilement de la menthe non traitée, contrairement à la menthe traitée avec Malyphos 50. Et sachant que l'ACP est une méthode linéaire, par conséquent elle n'est parfaitement pas adaptée aux réponses non linéaires de la matrice des capteurs de gaz. De plus, elle n'est pas dotée d'une règle de décision permettant d'identifier les échantillons inconnus. C'est pourquoi nous allons utiliser la méthode des machines à vecteurs de support (SVM).

I.3.2 Les machines à vecteurs de support (SVM)

La méthode des machines à vecteurs de support (SVM) est une méthode de classification supervisée, où le choix des groupes de données d'apprentissage se fait a priori [20], dans notre cas, nous nous sommes basés sur la nature des produits déjà connues. L'utilisation de la classification par des machines à vecteurs de support SVM vise à bien assigner les groupes de chaque classe de données. La classification a été effectuée à l'aide d'un classificateur SVM multi-classe utilisant une classification de partition Codage de sortie correcteur d'erreurs (ECOC, en anglais : Error-Correcting Output Coding) avec une validation croisée de 5-fold en utilisant les scores des individus obtenus à partir de l'analyse en composantes principales comme données de base, 80 % pour l'entraînement et 20 % pour le test. Le noyau de fonction de base radial a été utilisé pour projeter les données d'apprentissage dans un espace qui maximise la marge.

La méthode SVM a été appliquée aux scores issus de la méthode ACP relatifs aux différents types de menthe X1 et X2, X1 et X3, X2 et X3, et l'ensemble de données X4. Les résultats obtenus sont présentés dans la figure 4-3.

La figure 4-3 illustre les tracés des contours des probabilités postérieures sous les nuages de points des données. Les graphiques (a), (b) et (c) de la figure 4-3 ont montré les contours des régions des probabilités postérieures de chaque groupe de menthe avec une très bonne séparation tandis que le graphique (d) de la figure 4-3 a montré une bonne distinction entre les régions des probabilités postérieures de chaque type de menthe non traitée, traité avec le produit

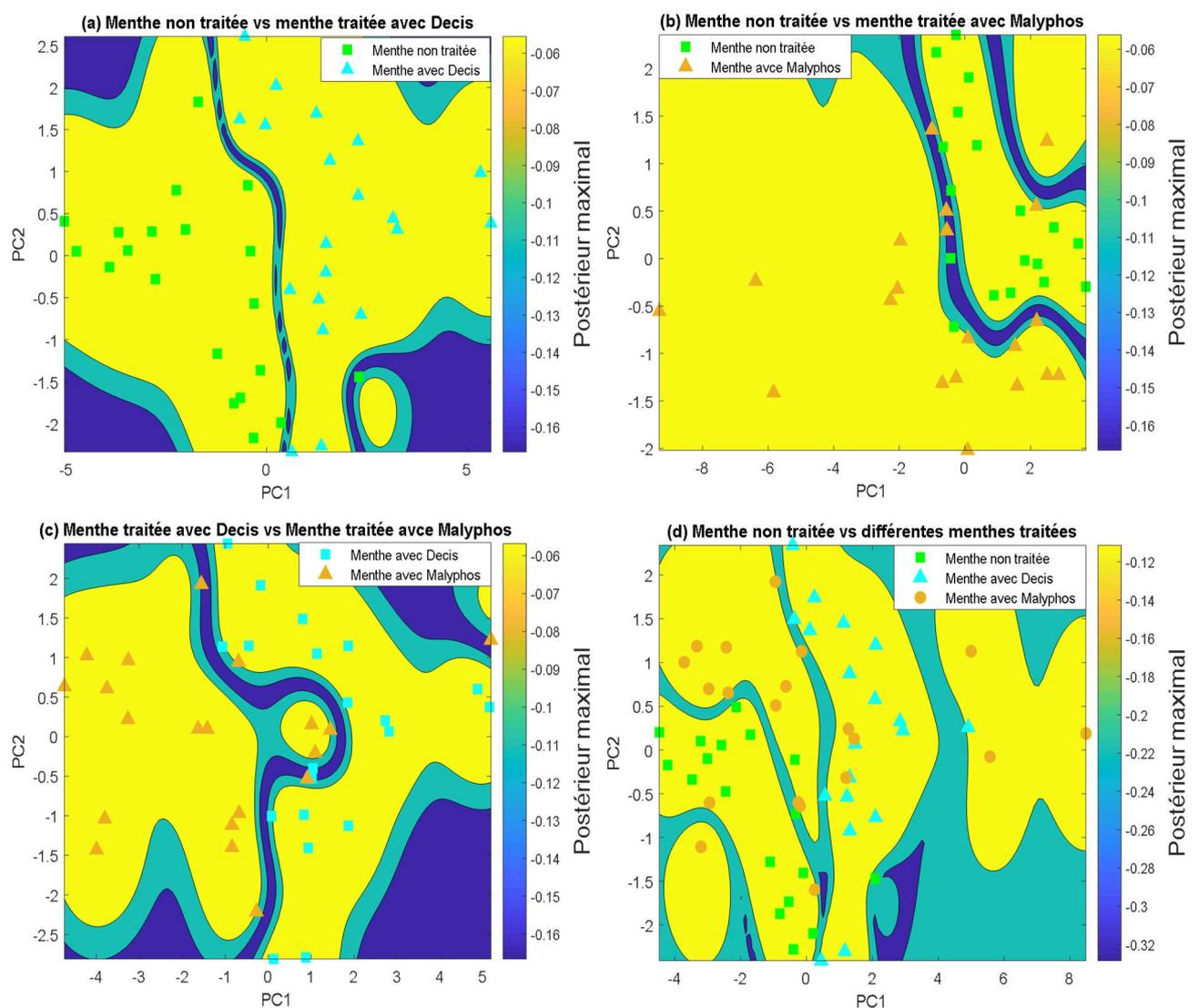


Figure 4-3 Résultats de la classification SVM des différents cas de menthe (a) menthe non traitée contre menthe traitée par Decis Fluxx, (b) menthe non traitée contre menthe traitée par Malyphos, (c) menthe traitée par Decis contre menthe traitée par Malyphos

Decis Fluxx et traité avec le produit Malyphos 50, ce qui était très difficile avec la méthode ACP.

À l'aide de cette méthode, nous avons obtenu les résultats représentés par les matrices de confusions présentes sur la figure 4-4.

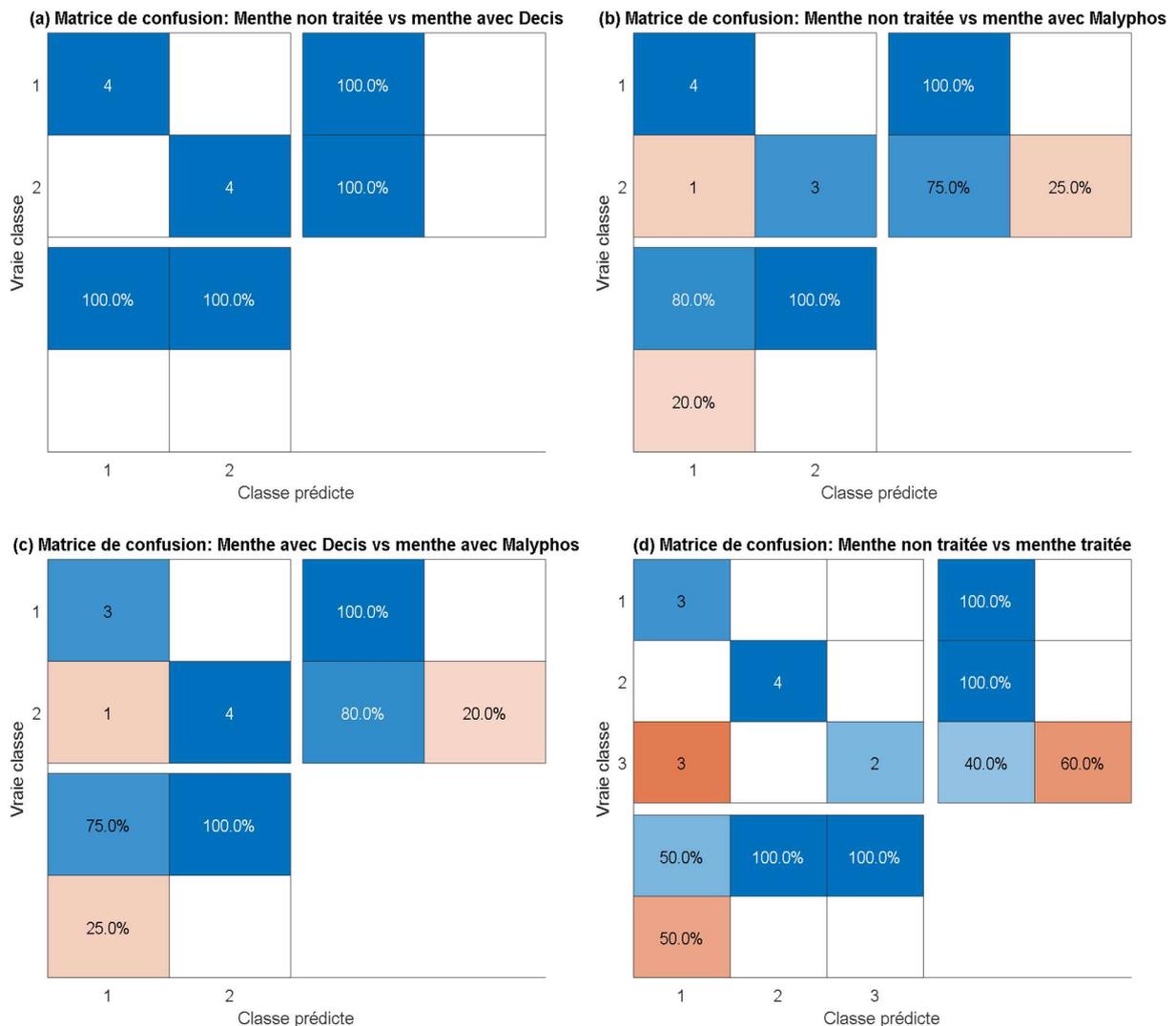


Figure 4-4 Matrice de confusion pour différents cas de menthe, (a) menthe non traitée contre menthe traitée par Decis Fluxx, (b) menthe non traitée contre menthe traitée par Malyphos, (c) menthe traitée par Decis contre menthe traitée par Malyphos, (d) menthe non traitée contre menthe traitée

Pour le premier cas de notre étude : la menthe non traitée contre la menthe traitée par Decis Fluxx, nous avons obtenu un taux d'identification de 100 %, dans un ensemble de 40 mesures (X1+X2) aucune erreur de classification n'a été enregistrée, toutes les données ont été bien classées.

Pour le deuxième et le troisième cas qui sont successivement le cas de la menthe non traitée contre la menthe traitée au Malyphos 50 et le cas de la menthe traitée au Decis Fluxx contre la menthe traitée au Malyphos 50, un taux d'identification de 87.5% a été réalisé. Dans des ensembles de quarante mesures, une seule erreur de classification sur huit échantillons de test a été enregistrée par cas d'étude : pour le deuxième cas (X1+X3), une mesure qui appartient a priori à la menthe traitée par Malyphos 50 a été mal classée dans la classe de la menthe non traitée, et pour le troisième cas (X2+X3), une mesure qui appartient a priori à la menthe traitée par Malyphos 50 a été mal classée dans la classe de la menthe traitée par Decis Fluxx.

Pour le quatrième et dernier cas : la menthe non traitée contre la menthe traitée, un taux d'identification de 75 % a été achevé, avec l'enregistrement de trois mesures sur douze échantillons de test d'un ensemble de soixante mesures (X4) qui ont été mal classées : les trois mesures qui appartiennent à la menthe traitée par Malyphos 50 ont été mal classées dans la classe de la menthe non traitée.

Les résultats montrent que les échantillons de test ont été affectés à leurs classes de données avec des taux de réussite encourageants, et selon ces résultats, nous pouvons dire que notre outil est capable de faire la distinction entre la menthe traitée et la menthe non traitée, en particulier celles traitées avec Decis Fluxx et peut être utilisé comme une autre alternative moins chère et facile à manipuler pour contrôler la qualité de la menthe.

I.4. Conclusion

Dans le cadre de cette étude, l'objectif était de distinguer entre trois types d'échantillons de la menthe comestible : menthe non traitée, menthe traitée avec l'insecticide Decis Fluxx et menthe traitée avec Malyphos 50. Nous avons prélevé des échantillons frais des trois classes préconnues, puis à l'aide d'un nez électronique conçu à base de capteurs de gaz à oxyde métallique de commerce, nous sommes parvenus à les bien trier selon leurs classes originales avec des taux de réussites très encourageants. En fait, L'utilisation de l'ACP en combinaison avec la méthode SVM a donné des taux de réussite entre 75% et 100% dans la discrimination des différents types de menthe. A l'issu des résultats de cette étude, nous pouvons dire qu'un simple, pas cher et portable nez électronique permet de distinguer la menthe non traitée de la menthe traitée avec des insecticides.

II. Deuxième Partie : ETUDE DE CAS DE TRAITEMENT DE LA MENTHE PAR LE MALATHION

Après la première étude menée sur les quatre premiers jours du premier traitement, qui a démontré l'utilité du nez électronique dans la discrimination de la menthe non traitée de celle traitée avec deux insecticides à savoir le Decis fluxx et le Malyphos 50 avec des taux de réussites très encourageants, cette partie vient pour approfondir l'étude sur la menthe traitée avec le Malyphos 50. Le malathion [21] est le composé principal de cet insecticide, il représente 50% de sa composition, ce composé est classé cancérigène par le Centre International de Recherche sur le Cancer (CIRC), et pour cette raison, son utilisation est interdite dans l'Union Européenne (décision 2007/389/CE).

L'identification de la présence des résidus de malathion, qui est un insecticide dangereux et cancérigène, dans les produits agricoles a fait l'objet de plusieurs études. Les techniques les plus couramment utilisées pour analyser le malathion sont les instruments spectrophotométriques et chromatographiques dont nous retrouvons notamment l'utilisation d'une spectrophotométrie cinétique catalytique [22] et la chromatographie en phase gazeuse [23] méthodes pour la détermination de la présence de résidus de malathion successivement dans des échantillons de légumes et de courgettes. La chromatographie liquide à haute performance-spectrométrie de masse en tandem (CLHP-SM/SM) pour la mesure des dégradants insecticides dont le malathion dans les aliments pour bébés [24]. L'extraction en phase solide et chromatographie en phase gazeuse avec spectrométrie de masse (EPS-CG-SM) pour la détermination du malathion dans différents produits alimentaires [25].

La présente étude a été réalisée dans l'objectif de proposer une résolution au problème du manque d'un outil simple, peu coûteux et portable pour le contrôle des insecticides neurotoxiques et cancérigènes comme le malathion dans la menthe audible.

II. 1. La préparation des échantillons

Le champ déjà cité dans la première partie de ce chapitre est celui qui est utilisé également dans cette deuxième. Comme nous avons vu et dans l'objectif d'être sûr des types de la menthe que nous allons utiliser, *Mentha Spicata L.* a été implantée à l'Ecole Nationale des Sciences Appliquées de Khouribga, Maroc (32°53 N 6°55 W) en espace libre d'une superficie de 9m². Après que la menthe ait bien mûri, nous avons gardé une première partie de menthe naturelle non traitée et la deuxième partie a été pulvérisée avec de l'insecticide Malyphos 50

contenant 50% de malathion qui est un composé organophosphoré cancérigène neurotoxique interdit dans plusieurs pays, sa formule chimique est $C_{10}H_{19}O_6PS_2$.

Nous envisagerons d'étudier le traitement de la menthe avec le produit Malyphos 50 pendant les sept jours du premier traitement puisque le producteur indique dans la notice que les produits agricoles traités avec son produit nécessitent sept jours de repos avant la récolte et la consommation, c'est pourquoi nous avons utilisé un litre de mélange pour pulvériser la partie consacrée à la menthe traitée comme indiqué sur la notice, ce mélange est composé de 4 ml de Malyphos 50 et d'un litre d'eau potable.

II.2. Réponses temporelles des capteurs

Les produits traités au malathion nécessitent un délai de sept jours avant la récolte et la consommation, c'est pourquoi nous avons décidé d'observer la menthe traitée avec ce produit sur les sept jours à compter du premier jour d'application de l'insecticide sur les zones du champ consacré à la menthe traitée. Dix expériences par jour divisées sur les deux types de menthe (cinq expériences par type) durant sept jours ont été faites comme suit : les échantillons avec un poids de $5g \pm 0,1g$ ont été prélevés frais et placés dans la chambre de l'échantillon, à partir de laquelle ils ont subi un flux d'air émis par le ventilateur pendant huit minutes avec un temps de 30 secondes avant de déclencher le flux d'air pour observer la stabilité des capteurs.

Les courbes des réponses en tension relative des capteurs en fonction du temps enregistrées du premier jour du traitement au septième jour sont présentées par la figure 4-5.

Selon les résultats, les tensions délivrées par les capteurs ont augmenté en présence de la menthe traitée avec du malathion, prouvant ainsi la capacité des capteurs à détecter l'évolution de la composition de l'odeur de menthe.

En présence des composants volatils de l'espace d'échantillon de la menthe, un changement de tension important est mesuré en sortie des capteurs, avec une différence dans les réponses suivant le jour de traitement.

En observant les deux graphes (a) et (b) correspondant consécutivement aux capteurs TGS-822 et TGS-2620 sur la figure 4-5. On remarque que les réponses relatives à la menthe traitée dans les sept jours sont supérieures à la menthe non traitée, par ailleurs, on remarque que les réponses sont progressives du premier jour au quatrième jour qu'elles atteignent le maximum puis elles déminent du quatrième au septième jour pour le cas des réponses du capteur TGS-

822. Même chose pour le cas du capteur TGS-2620 avec une seule différence, le maximum est atteint au cinquième jour. Cela peut être analysé comme suit : pour les quatre ou cinq premiers jours, l'augmentation de la concentration des composés chimiques des composés volatils est certainement dû à l'apparition des nouvelles substances chimiques à lesquelles la matrice des capteurs est très sensible. Après cet intervalle de temps, la concentration des composés chimiques qui sont apparus décroît progressivement pour revenir à l'état de menthe non traitée.

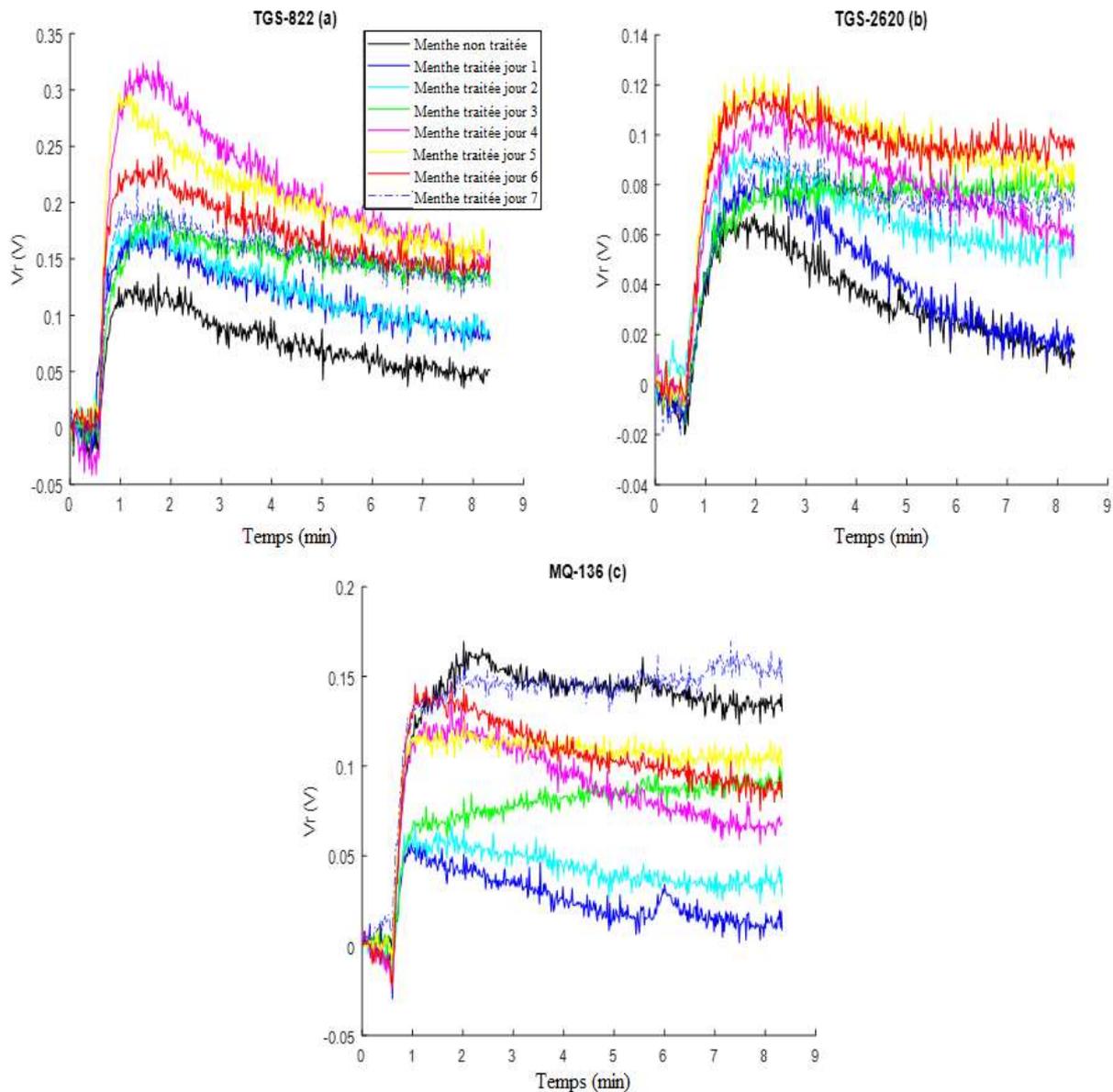


Figure 4-5 Réponses temporelles des trois capteurs de gaz pour la menthe non traitée et la menthe traitée en fonction des jours

En examinant le graphe (c) correspondant au MQ-136 sur la figure 4-5, on constate que l'intensité du signal relatif à la menthe non traitée est la plus grande et elle est

approximativement équivalente à celle du signal relatif au septième jour de la menthe traitée. Une autre remarque très importante, c'est que les intensités des signaux de la menthe traitée s'accroissent progressivement d'un jour à l'autre du premier jour au septième où le signal du septième jour est presque équivalent au signal de la menthe non traitée comme signalé précédemment. Cela peut s'expliquer de la manière suivante : l'ajout du produit du traitement (l'insecticide) a modifié la composition des composés volatils générés par la menthe de telle façon à ce que les capteurs de la matrice n'arrivaient pas à bien les détecter, mais avec le temps, la menthe parvient à retrouver son arôme initial.

Les informations extraites des réponses des capteurs vont nous aider à faire la distinction entre menthe non traitée et menthe traitée et même de définir le jour de traitement de la menthe en utilisant la sensibilité croisée.

II.3. Discrimination de la menthe traitée de celle non traitée

Après l'étape du prétraitement des données, les caractéristiques extraites des signaux ont été organisées en deux ensembles de données :

- X1 : ensemble de données totales contenant toutes les données recueillies relatives aux données de la menthe non traitée et celle traitée au Malyphos 50. Il est composé de neuf colonnes (3 réponses des capteurs * 3 caractéristiques principales (valeur max + l'aire du signal + la valeur stabilisée)) et de soixante-dix lignes (5 échantillons * 2 types de menthe * 7 nombres de jours).
- X2 : ensemble de données relatives à la menthe traitée au Malyphos 50 uniquement. Il est composé de neuf colonnes (3 réponses des capteurs * 3 caractéristiques principales) et de trente-cinq lignes (5 échantillons * 1 type de menthe * 7 nombres de jours).

II.3.1. Projection de différents types de menthe par ACP à trois dimensions

La projection de l'ensemble de données X1 regroupant les données de la menthe traitée et celle non traitée par PCA est illustrée à la figure 4-6. Dans notre cas, elle est utilisée pour réduire notre espace de données à neuf dimensions à un espace de données à trois dimensions (PC1, PC2 et PC3).

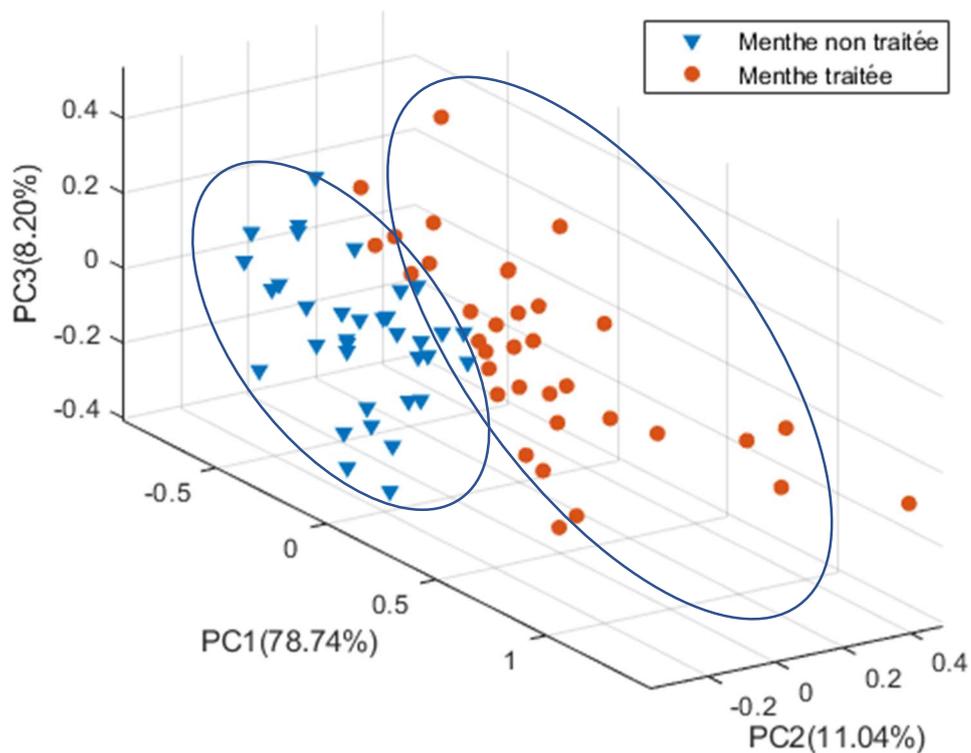


Figure 4-6 Projection des données sur les trois premières composantes principales de l'ACP

Le taux de contribution de la première composante principale (PC1) était de 78,74 %, de la deuxième composante principale (PC2) était de 11,04 %, de la troisième composante principale (PC3) était de 8,20 %. Le taux de contribution cumulé des trois premières composantes principales était de 97,98 %, ce qui signifie qu'un très bon taux d'informations contenues dans l'ensemble de données a été retenu et représenté. Etant donné que l'ACP aide à regrouper les données ayant des propriétés similaires, le nuage de points en trois dimensions a montré que les échantillons d'un même groupe sont corrélés et regroupés et les deux groupes se chevauchent légèrement, ce qui permet de les identifier visuellement.

Le résultat obtenu peut être exploité pour les échantillons de la menthe inconnus après avoir construit le modèle grâce aux deux groupes de données. Pour évaluer la qualité d'un échantillon de menthe (traitée ou non), on expose la matrice des capteurs à un flux d'air enrichi des composés organiques volatils de ce dernier, et les valeurs extraites des signaux de réponses sont ajoutées à l'ancienne matrice de données. Le lieu de projection de l'individu par le modèle de classification déterminera alors à quel groupe appartient l'échantillon.

Ce résultat prometteur de la méthode ACP montre qu'il est possible de distinguer visuellement les différents types de menthe. Sachant que cette méthode ne donne que des représentations graphiques des données en classes et elle ne dispose pas de règle de décision permettant d'identifier les échantillons inconnus [26], nous utiliserons par la suite la méthode k-plus proche voisin (KNN). Généralement, on s'attend à ce que les ensembles des données qui présentent un léger chevauchement donnent de bonnes performances de classification.

II.3.2. Identification de la menthe non traitée à partir de la menthe traitée par la méthode KNN

Comme décrite dans le chapitre II parmi les algorithmes de l'apprentissage automatique des données, la méthode KNN est une méthode supervisée, donc le choix des groupes de données d'apprentissage se fait a priori, dans notre cas, nous nous sommes basés sur la nature de la menthe déjà connue (traitée ou non traitée). Cette méthode repose principalement sur la définition de deux paramètres principaux : K le nombre de voisins et la métrique de distance. En évaluant la capacité de prédiction de plusieurs valeurs de K, nous avons découvert que K = 5 est le meilleur choix pour notre application, car il a donné le plus grand taux de classification. Quant à la métrique de distance choisie, elle est euclidienne.

Pour évaluer le modèle de prédiction, nous avons utilisé la validation croisée 5-fold qui consiste à diviser l'ensemble de données en cinq sous-groupes de données : quatre groupes seront utilisés comme une base d'apprentissage et le dernier groupe sera utilisé pour le test. Cette approche a été refaite à cinq reprises en faisant passer les cinq sous-groupes de l'apprentissage et de prédiction, garantissant de ce fait que tous les échantillons puissent participer dans la base de test. Le taux de classification global est défini par le taux moyen des cinq tests. L'analyse et le traitement des données ont été effectués à l'aide de la boîte à outils de statistiques et d'apprentissage automatique du logiciel MATLAB (2019a).

Pour visualiser le résultat, nous avons choisi la valeur maximale extraite du capteur MQ136 et celle extraite du capteur TGS822 pour visualiser les résultats dans un plan bidimensionnel comme illustré sur la figure 4-7.

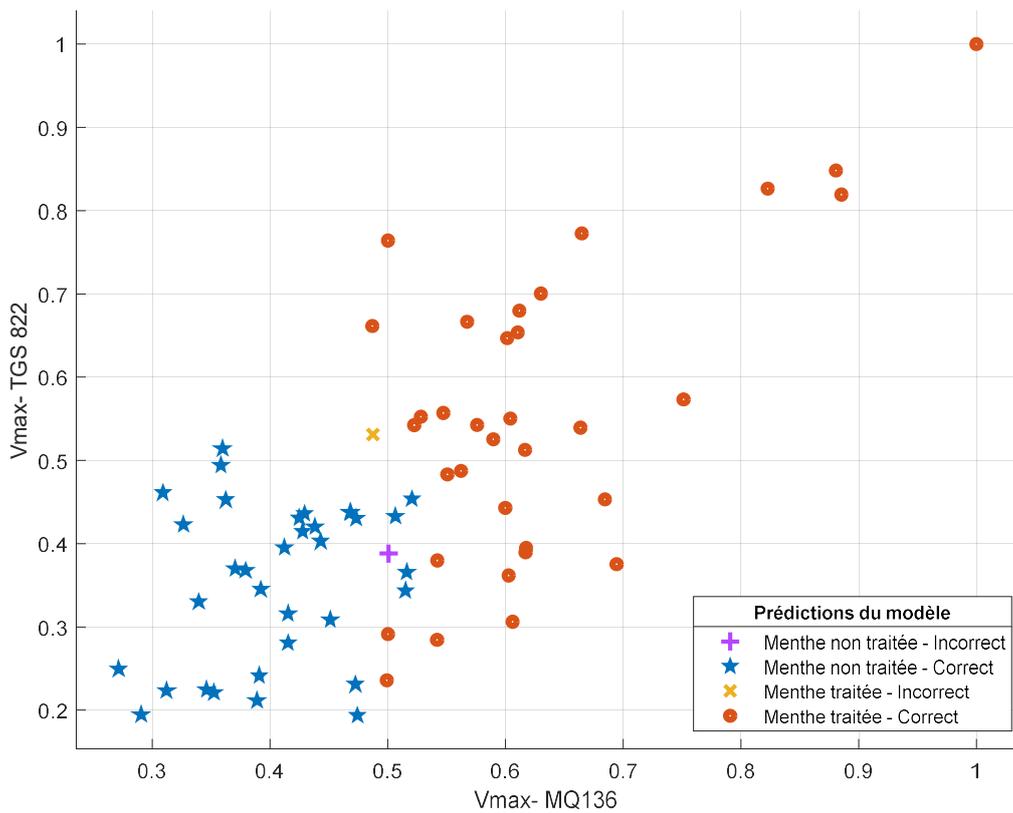


Figure 4-7 Résultat de KNN pour l'identification du type de menthe

L'algorithme choisi est arrivé à bien classer les groupes des données. Le taux de réussite pour le premier groupe relatif à la menthe non traitée est de 97%, le même pour le deuxième groupe relatif à la menthe traitée (figure 4-8).

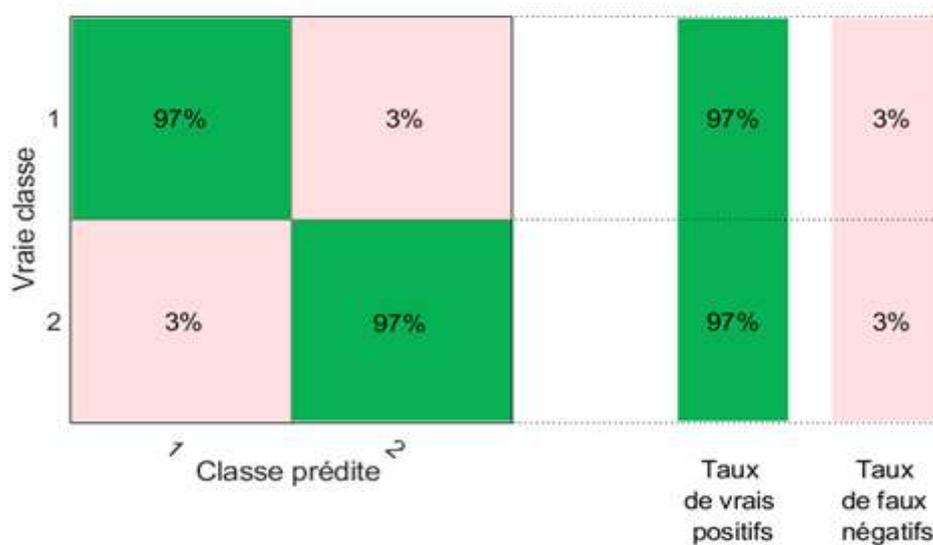


Figure 4-8 Taux de réussite de KNN pour l'identification du type de menthe

Les échantillons ont été correctement affectés à leurs classes de données, à l'exception d'un échantillon de chaque type de menthe qui a été mal classé, mais malgré cela, le taux de réussite était excellent, environ 97,10 %, un très bon résultat par rapport à l'étude précédente [27] qui a été menée sur les quatre jours seulement après la pulvérisation de l'insecticide contenant du Malathion, et qui a eu un taux de réussite de 87.5% seulement. Il est à noter que dans la précédente étude, une méthode hybride a été utilisée qui consiste à introduire les scores de l'analyse en composantes principales (ACP) comme données de base pour l'algorithme de séparateur à vaste marge (SVM).

Sur la base de ces bons résultats de classification, la méthode simple KNN peut être considérée comme la stratégie la plus puissante pour créer des règles de décision sûres avec un pourcentage de classification élevé. Ce qui nous permet de dire que notre outil couplé à la méthode d'apprentissage automatique KNN est capable de faire la distinction entre la menthe traitée au malathion et la menthe non traitée pendant les sept jours après la première application de l'insecticide.

II.4. Prédiction du jour de traitement de la menthe avec le réseau NARX

L'idée de cette étude vient du fait que le producteur du produit Malyphos 50 déclare dans la notice que les produits agricoles traités avec son produit nécessitent sept jours de repos avant la récolte et la consommation. Comme évoqué précédemment dans le paragraphe de l'analyse des réponses des capteurs dans la section des réponses temporelles, nous avons observé ce qui suit : pour les deux capteurs TGS822 et TGS2620, nous avons remarqué que les réponses sont progressives du premier au quatrième jour (pour TGS2620) ou au cinquième jour (pour TGS822) dont elles atteignent la maximale valeur puis elles diminuent. Pour le capteur MQ136, les intensités des signaux de la menthe traitée s'accroissent progressivement d'un jour à l'autre du premier jour au septième où le signal du septième jour est presque équivalent au signal de la menthe non traitée. Ces remarques qu'on peut les interpréter par une forte relation entre les réponses des capteurs et le jour du prélèvement de la menthe traitée, d'où l'idée d'utiliser les données des trois capteurs avec une méthode de régression pour prédire le jour du prélèvement de l'échantillon de la menthe traité avec du malathion.

Pour prédire le jour du prélèvement de la menthe traitée en se servant des composés volatils engendrés par les échantillons, l'ensemble de données X2 relatif à la menthe traitée et un réseau NARX sont utilisés. Dans notre cas, nous avons un ensemble de données de 35 lignes

et 9 colonnes, elles ont été allouées comme suit 70% pour l'entraînement, 15% pour la validation et 15% pour le test. Pour la rétropropagation, L'algorithme de régularisation bayésienne a été utilisé, c'est l'algorithme d'entraînement le plus adapté pour le traitement de petits ensembles de données, il permet d'obtenir la meilleure solution même si cela demande plus de temps pour l'exécution.

Après plusieurs tests, le réseau a été testé en faisant varier le nombre de neurones dans la couche cachée et le délai des périodes pour trouver la meilleure combinaison entre ces paramètres. Un modèle avec dix couches cachées et un délai de trois a été retenu. La figure 4-9 montre le schéma fonctionnel du réseau NARX utilisé dans MATLAB.

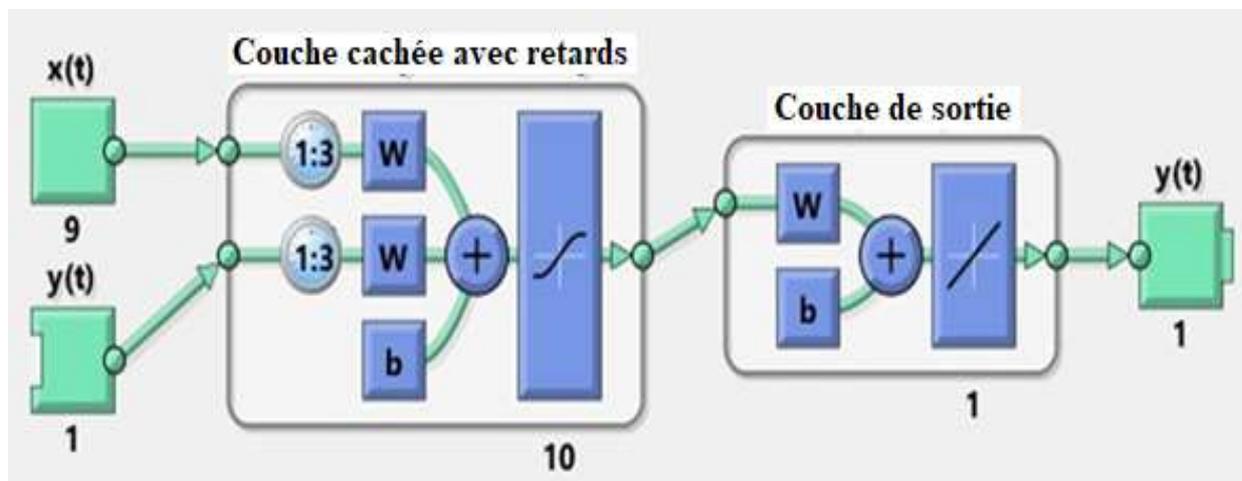


Figure 4-9 Architecture du réseau NARX dans Matlab

Après l'exécution du programme et d'après la figure 4-10, dans un premier temps, Nous observons une diminution de l'erreur quadratique moyenne sur les deux ensembles de données de l'entraînement et du test, puis après un certain temps nous remarquons que l'erreur quadratique moyenne commence légèrement à augmenter sur l'ensemble de test au même temps qu'elle continue de décroître sur l'ensemble d'apprentissage. L'apprentissage s'arrête dès que les meilleures performances sont atteintes (indiqués par la ligne pointillée)

L'entraînement a été arrêté après 196 itérations lorsque le gradient a atteint son minimum. L'erreur quadratique moyenne est très petite est d'environ $1,2986e-14$.

Meilleure performance d'entraînement est 1.2986e-14 à l'itération 196

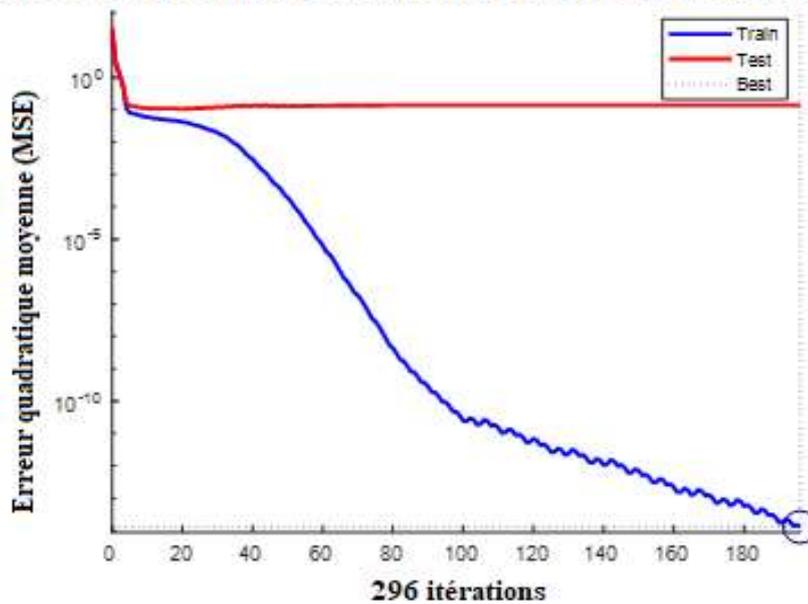


Figure 4-10 Performance du réseau NARX

Les diagrammes des régressions, qui affichent les régressions linéaires entre les objectifs ciblés et les sorties de réseau associées (sorties prévues), sont des autres vérifications cruciales des performances des réseaux. Les sorties et les objectifs doivent être fixés sur la ligne à 45 degrés et, plus simplement, précisément égaux avec une petite imprécision dans des conditions d'entraînement parfaites. La figure 4-11 illustre le graphique de régression de cette étude. L'ajustement entre les données cibles souhaités et la sortie du réseau NARX semble bien bon pour l'entraînement, le test et l'ensemble de données total avec l'enregistrement d'un coefficient de corrélation R idéal ($R = 1$) pour l'entraînement (Fig. 4-11 (a)), pour le test, il est de 0,981 (Fig. 4-11 (b)), et pour l'ensemble de données total, il est approximativement idéal est de 0,997 proche de 1 (Fig. 4-11 (c)). Par conséquent, il existe une bonne approximation entre les valeurs expérimentales originales et les sorties prévues du réseau.

Sur la base des indicateurs de performance, l'erreur quadratique moyenne (MSE) et le coefficient de corrélation (R), le jour de traitement de la menthe au malathion en fonction des modifications des composés volatils peut être estimée avec une grande précision.

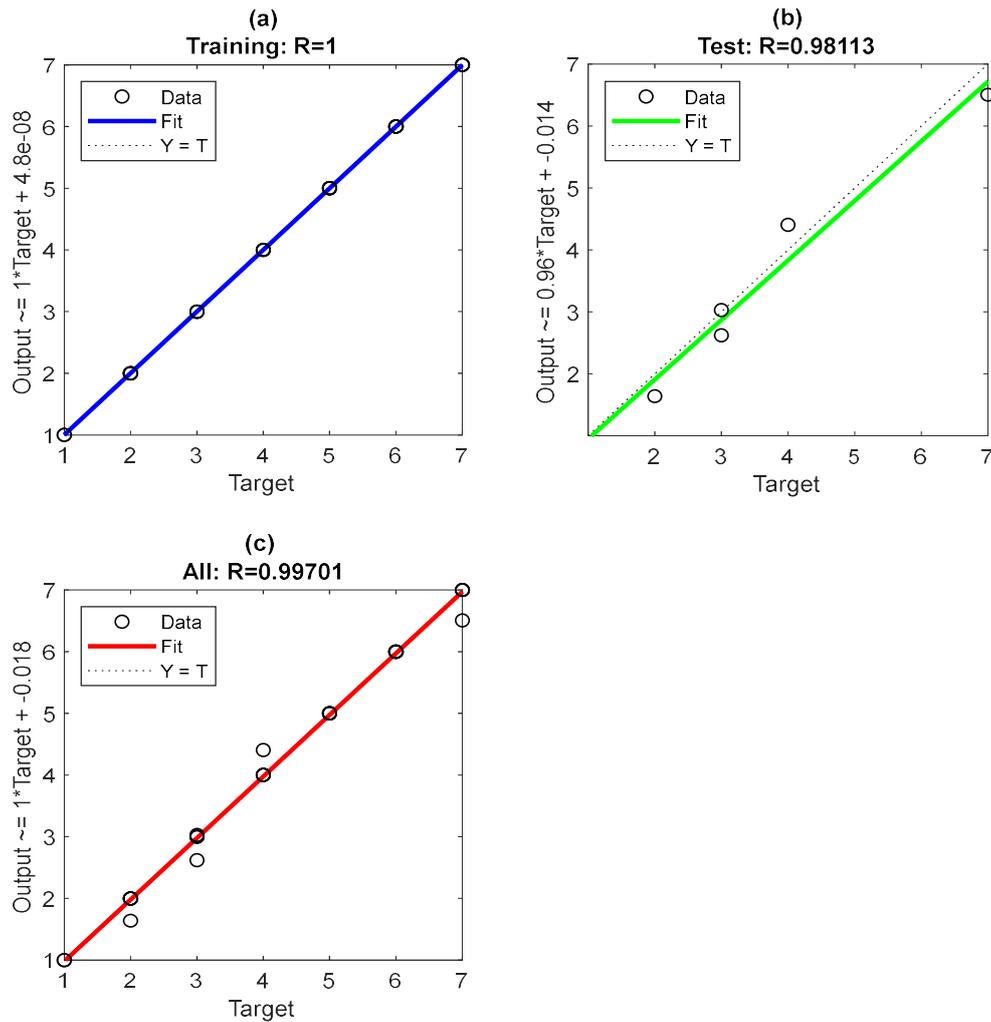


Figure 4-11 Corrélation entre les sorties et les valeurs cibles. (a) pour les données d'apprentissage, (b) pour le test et (c) pour toutes les données

II.5. Comparaison des résultats du réseau NARX avec PLSR et ϵ -SVM

Dans le cadre de la valorisation des résultats obtenus par le réseau NARX, une comparaison à d'autres algorithmes de régression à savoir la régression par moindres carrés partielles PLSR et la régression linéaire des séparateurs à vaste marge ϵ -SVM a été réalisée.

La régression PLSR est un algorithme qui crée un modèle linéaire prédictif qui représente les variables d'entrée X (les caractéristiques extraites des signaux des capteurs) en fonction d'une variable de réponse Y (le jour du prélèvement de la menthe traitée). Cette régression se réalise par l'extraction des variables latentes dérivées de la matrice des variables prédictives X. Le PLSR crée ces variables latentes en considérant la corrélation entre les variables X ainsi que la covariance entre les variables X et Y. Les premières variables latentes

créées par PLSR sont connectées aux informations contenues dans les données, tandis que les dernières sont associées au bruit, de la même manière que les principales composantes sont calculées dans l'algorithme de l'ACP.

Le choix du nombre pertinent des variables latentes qui doivent être prises en considération est le critère qui influence l'efficacité du modèle. En effet, ce nombre doit être minimale pour prévenir le sur-apprentissage mais suffisamment grand pour empêcher le sous-apprentissage.

Dans ce présent cas, les trois premières variables latentes ont été choisies et nous avons utilisé la validation hold-out, 80% des données pour l'apprentissage et 20% pour le test. La figure 4-12 représente la droite de régression entre les réels jours du prélèvement de la menthe traitée et les jours prédits par le modèle PLSR pour les données d'entraînement et de test. Nous apercevons que les différentes observations sont loin de la droite de régression, ce qui prouve une mauvaise corrélation entre les valeurs observées et les valeurs prédites. Donc la régression PLSR n'a pas réussi à bien prédire les jours avec une bonne précision, le coefficient de corrélation est faible de 0,67 et l'erreur quadratique moyenne (MSE) est très grande de 2,0958.

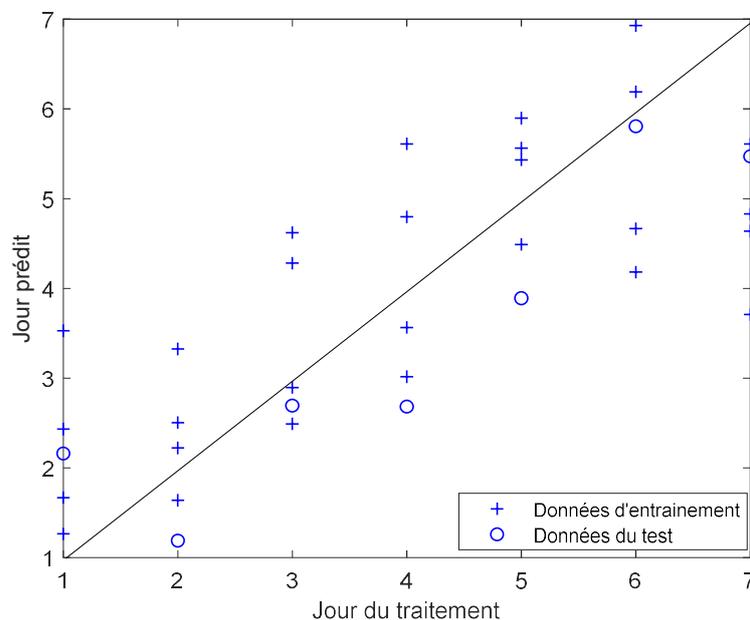


Figure 4-12 Résultat de corrélation de l'algorithme PLSR pour la prédiction du jour du prélèvement de la menthe traitée

Pour la régression ϵ -SVM, c'est une régression qui cherche à trouver une transformation diagonale linéaire et une translation qui maximisent la marge par rapport à l'hyperplan optimal

en maintenant tous les individus proches de ce dernier avec la condition : chaque point d'apprentissage ne s'écarte de l'hyperplan optimal que d'une valeur ne dépassant pas pour la valeur de ϵ .

Pour la régression ϵ -SVM, la validation croisée 5-folds a été utilisée. La figure 4-13 représente la droite de régression entre les réels jours du prélèvement de la menthe traitée et les jours prédits par la régression ϵ -SVM. Nous constatons que les différents individus sont moins loin de la droite de régression, un résultat mieux que celui réalisé par la régression PLSR. Mais ça n'empêche pas de dire qu'il y a toujours une mauvaise corrélation entre les valeurs observées et les valeurs prédites. Également, la régression ϵ -SVM n'a pas réussi à bien prédire les jours avec une bonne précision, le coefficient de corrélation est moins bon de 0,82 et l'erreur quadratique moyenne (MSE) est relativement élevée de 1,3615.

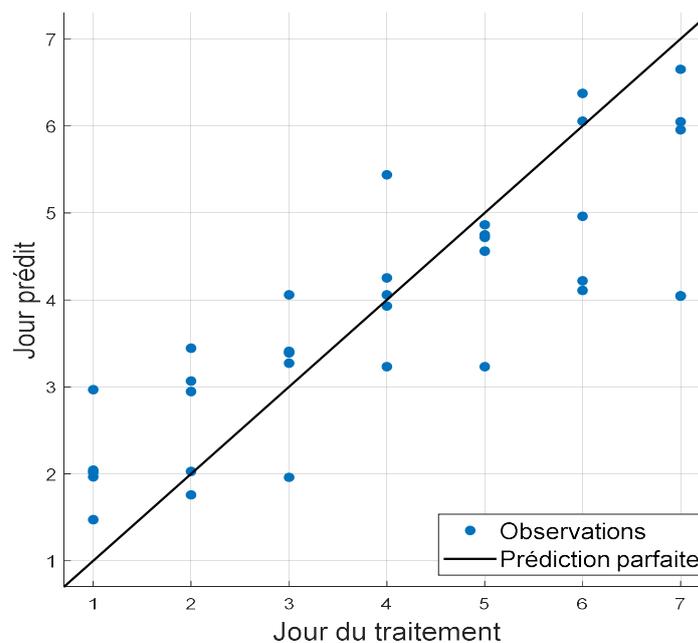


Figure 4-13 Résultat de corrélation de l'algorithme ϵ -SVM pour la prédiction du jour du prélèvement de la menthe traitée

En ajoutant les résultats obtenus avec le réseau NARX, L'étude comparative pour l'évaluation de l'efficacité des algorithmes proposés pour la prédiction du jour de traitement de la menthe avec du malathion est illustrée dans le tableau 4-1.

Tableau 4-1 Résultats de l'étude comparative des méthodes de régression

METHODE	Caractéristiques	
	Corrélation coefficient (R)	Erreur quadratique moyenne (MSE)
Régression des moindres carrés partiels (PLSR)	0.67	2.0958
Régression ϵ -SVM	0.82	1.3615
Réseau NARX	0.99	1.10288e-14

Dans l'ensemble, le réseau NARX a obtenu le meilleur résultat, le coefficient de corrélation est de 0,99 et l'erreur quadratique moyenne (MSE) est de 1,10288e-14 très proche de zéro.

Ces résultats ont montré clairement la supériorité du réseau NARX dans notre cas pour la prédiction du jour de traitement de la menthe au malathion.

Pour conclure, le système multi-capteurs (nez électronique) peut être utilisé comme un dispositif utile pour l'identification de la menthe traitée avec malathion ainsi que son jour de traitement en fonction des composés volatils dégagés par des échantillons.

II.4. Conclusion

L'utilisation intensive d'insecticides dans l'agriculture nécessite une grande vigilance en raison des effets indésirables qui mettent la vie humaine en danger. Dans le présent travail, un système multi-capteurs (nez électronique) fabriqué basé sur trois capteurs de gaz commerciaux a été utilisé pour analyser les composés volatils de la menthe pendant les sept jours de traitement avec l'insecticide malathion. Dans la première étape, il a été démontré qu'en utilisant trois caractéristiques de réponse des capteurs qu'il est possible d'identifier la menthe traitée et celle non traitée en utilisant la méthode des k-plus proches voisins (KNN) avec la validation croisée 5 fois qui a donné les résultats 97% de taux de réussite pour l'identification de la nature de la menthe. Dans la deuxième étape, il a été validé qu'il est possible de prédire le jour de traitement de l'échantillon de la menthe en utilisant le réseau autorégressif non linéaire avec entrée exogène

(NARX) en réalisant un coefficient de corrélation de 0,99 et une erreur quadratique moyenne presque nulle. À l'issue de ce travail, on peut dire qu'il est possible d'utiliser un simple et portable nez électronique à faible coût pour discriminer entre la menthe traitée et non traitée et prédire avec une grande précision le jour de traitement de la menthe.

Références

- [1] J. R. Hayes, M. S. Stavanja, and B. M. Lawrence, Biological and toxicological properties of mint oils and their major isolates: Safety assessment. 2006.
- [2] T. Eddaya, A. Boughdad, L. Becker, P. Chaimbault, and A. Zaïd, "Utilisation et risques des pesticides en protection sanitaire de la menthe verte dans le Centre-Sud du Maroc," *J. Mater. Environ. Sci.*, vol. 6, no. 3, pp. 656–665, 2015.
- [3] A. Kalra, H. B. Singh, R. Pandey, A. Samad, N. K. Patra, and S. Kumar, "Diseases in mint: Causal organisms, distribution, and control measures," *J. Herbs, Spices Med. Plants*, vol. 11, no. 1–2, pp. 71–91, 2004, doi: 10.1300/J044v11n01_03.
- [4] P. Nicolopoulou-Stamati, S. Maipas, C. Kotampasi, P. Stamatis, and L. Hens, "Chemical Pesticides and Human Health: The Urgent Need for a New Concept in Agriculture," *Front. Public Heal.*, vol. 4, no. July, pp. 1–8, 2016, doi: 10.3389/fpubh.2016.00148.
- [5] W. Dong, Y. Ni, and S. Kokot, "Differentiation of Mint (*Mentha haplocalyx* Briq.) from different regions in China using gas and liquid chromatography," *J. Sep. Sci.*, vol. 38, no. 3, pp. 402–409, 2015, doi: 10.1002/jssc.201401130.
- [6] Y. Yu et al., "Successive monitoring surveys of selected banned and restricted pesticide residues in vegetables from the northwest region of China from 2011 to 2013," *BMC Public Health*, vol. 18, no. 1, pp. 1–9, 2017, doi: 10.1186/s12889-017-4632-x.
- [7] A. Diop et al., "Monitoring survey of the use patterns and pesticide residues on vegetables in the Niayes zone, Senegal," *Chemosphere*, vol. 144, pp. 1715–1721, 2016, doi: 10.1016/j.chemosphere.2015.10.058.
- [8] M. F. A. Jallow, D. G. Awadh, M. S. Albaho, V. Y. Devi, and N. Ahmad, "Monitoring of pesticide residues in commonly used fruits and vegetables in Kuwait," *Int. J. Environ. Res. Public Health*, vol. 14, no. 8, 2017, doi: 10.3390/ijerph14080833.
- [9] J. H. Y. Galani, M. Houbraken, A. Wumbei, J. F. Djeugap, D. Fotio, and P. Spanoghe, "Evaluation of 99 pesticide residues in major agricultural products from the Western Highlands zone of Cameroon using QuEChERS method extraction and LC-MS/MS and GC-ECD analyses," *Foods*, vol. 7, no. 11, pp. 1–17, 2018, doi: 10.3390/foods7110184.
- [10] M. A. Farajzadeh, M. R. Afshar Mogaddam, S. Rezaee Aghdam, N. Nouri, and M. Bamorowat, "Application of elevated temperature-dispersive liquid-liquid microextraction for determination of organophosphorus pesticides residues in aqueous samples followed by gas chromatography-flame ionization detection," *Food Chem.*, vol. 212, pp. 198–204, 2016, doi: 10.1016/j.foodchem.2016.05.157.
- [11] L. Baraud, D. Tessier, J. J. Aaron, J. P. Quisefit, and J. Pinart, "A multi-residue method

- for characterization and determination of atmospheric pesticides measured at two French urban and rural sampling sites,” *Anal. Bioanal. Chem.*, vol. 377, no. 7–8, pp. 1148–1152, 2003, doi: 10.1007/s00216-003-2196-3.
- [12] M. Pirsaeheb, N. Fattahi, M. Karami, and H. R. Ghaffari, “Simultaneous determination of deltamethrin, permethrin and malathion in stored wheat samples using continuous sample drop flow microextraction followed by HPLC–UV,” *J. Food Meas. Charact.*, vol. 12, no. 1, pp. 118–127, 2018, doi: 10.1007/s11694-017-9622-2.
- [13] G. Kowalska, U. Pankiewicz, and R. Kowalski, “Estimation of pesticide residues in selected products of plant origin from Poland with the use of the HPLC-MS/MS technique,” *Agric.*, vol. 10, no. 6, pp. 1–20, 2020, doi: 10.3390/agriculture10060192.
- [14] Y. Huang et al., “Determination of multi-pesticide residues in green tea with a modified QuEChERS protocol coupled to HPLC-MS/MS,” *Food Chem.*, vol. 275, no. August 2018, pp. 255–264, 2019, doi: 10.1016/j.foodchem.2018.09.094.
- [15] H. El-Haoud, M. Boufellous, M. Faid, and R. Bengueddour, “Risks of Phytosanitary Uses and Residue Analysis Pesticides on Green Mint (*Mentha Spicata* L.) in the Province of Benslimane in Morocco,” *Am. J. Innov. Res. Appl. Sci.*, no. December 2017, pp. 2429–5396, 2017, [Online]. Available: www.american-jiras.com.
- [16] R. Mestres and G. Mestres, “Deltamethrin: Uses and environmental safety,” *Rev. Environ. Contam. Toxicol.*, vol. 124, pp. 1–18, 1992, doi: 10.1007/978-1-4612-2864-6_1.
- [17] P. Flessel, P. J. E. Quintana, and K. Hooper, “Genetic toxicity of malathion: A review,” *Environ. Mol. Mutagen.*, vol. 22, no. 1, pp. 7–17, 1993, doi: 10.1002/em.2850220104.
- [18] S. Bedoui, R. Faleh, H. Samet, and A. Kachouri, “Electronic nose system and principal component analysis technique for gases identification,” 2013 10th Int. Multi-Conference Syst. Signals Devices, SSD 2013, pp. 1–6, 2013, doi: 10.1109/SSD.2013.6564152.
- [19] W. Jia, G. Liang, Z. Jiang, and J. Wang, “Advances in Electronic Nose Development for Application to Agricultural Products,” *Food Anal. Methods*, vol. 12, no. 10, pp. 2226–2240, 2019, doi: 10.1007/s12161-019-01552-1.
- [20] T. Oommen, D. Misra, N. K. C. Twarakavi, A. Prakash, B. Sahoo, and S. Bandopadhyay, “An objective analysis of support vector machine-based classification for remote sensing,” *Math. Geosci.*, vol. 40, no. 4, pp. 409–424, 2008, doi: 10.1007/s11004-008-9156-6.
- [21] W. N. Aldridge, J. W. Miles, D. L. Mount, and R. D. Verschoyle, “The toxicological properties of impurities in malathion,” *Arch. Toxicol.*, vol. 42, no. 2, pp. 95–106, 1978, doi: 10.1007/BF00316489.
- [22] N. Tiwari and A. Asthana, “A catalytic kinetic spectrophotometric determination of organophosphorus pesticides in vegetable samples,” *J. Braz. Chem. Soc.*, vol. 23, no. 2, pp. 322–327, 2012, doi: 10.1590/S0103-50532012000200019.
- [23] H. M. Lofty, A. E.-A. A. Abd El-Aleem, and H. H. Monir, “Determination of insecticides malathion and lambda-cyhalothrin residues in zucchini by gas

- chromatography,” *Bull. Fac. Pharmacy, Cairo Univ.*, vol. 51, no. 2, pp. 255–260, 2013, doi: 10.1016/j.bfopcu.2013.08.001.
- [24] S. A. Radford, P. Panuwet, R. E. Hunter, D. B. Barr, and P. B. Ryan, “HPLC-MS/MS method for the measurement of insecticide degradates in baby food,” *J. Agric. Food Chem.*, vol. 62, no. 29, pp. 7085–7091, 2014, doi: 10.1021/jf500779a.
- [25] A. Vijaya Bhaskar Reddy et al., “Simultaneous determination of three organophosphorus pesticides in different food commodities by gas chromatography with mass spectrometry,” *J. Sep. Sci.*, vol. 39, no. 12, pp. 2276–2283, 2016, doi: 10.1002/jssc.201600155.
- [26] A. W. Whitney, “A Direct Method of Nonparametric Measurement Selection,” *IEEE Trans. Comput.*, vol. C–20, no. 9, pp. 1100–1103, 1971, doi: 10.1109/T-C.1971.223410.
- [27] A. Amkor and N. El Barbri, “A measurement prototype based on gas sensors for detection of pesticide residues in edible mint,” *J. Food Meas. Charact.*, 2020, doi: 10.1007/s11694-020-00617-8.

**Chapitre 5 : RESULTATS ET
DISCUSSIONS : ETUDE DE CAS DES
POMMES DE TERRE**

Introduction

Après le blé, le riz et le maïs, la pomme de terre (*Solanum tuberosum* L.) est le quatrième produit agricole le plus important au monde, elle est cultivée dans la plupart des zones climatiques occupées par l'homme [1]. Ce produit végétal originaire d'Amérique du Sud est largement consommé en raison de sa valeur nutritionnelle, il est très riche dans sa composition en amidon qui est un glucide complexe, des protéines brutes, une variété de vitamines et minéraux essentiels dont les vitamines C et B6, et les minéraux potassium, magnésium et fer, en plus des fibres alimentaires [2].

Le présent chapitre présentera les résultats des travaux de recherche en utilisant l'outil du nez électronique conçu et réalisé au sein du laboratoire pour distinguer entre les pommes de terre en se basant sur leurs provenances, pommes de terre récoltées à partir d'un champ traité traditionnellement avec le fumier des moutons et des ânes, champ traité avec les engrais NPK (N : azote, P : phosphore et K : potassium) et champ traité avec le fumier de poulets. Ce chapitre est organisé en deux parties distinctes, dans la première nous allons montrer que notre système est capable de distinguer entre les pommes de terre traitées traditionnellement et celles traitées avec les engrais NPK dans un premier temps. Par la suite et dans la deuxième partie, nous allons étudier les trois types de pommes de terre : traitées traditionnellement, traitées avec les engrais NPK et traitées avec le fumier de poulet, et nous montrerons que notre nez électronique est toujours capable de distinguer entre les trois types précités.

I. Première Partie : IDENTIFICATION DES POMMES DE TERRE TRAITÉES AVEC LES ENGRAIS NPK DE CELLES TRAITÉES TRADITIONNELLEMENT

La demande croissante, ainsi que la consommation importante de pommes de terre, poussent les agriculteurs à utiliser des engrais biologiques ou chimiques pour améliorer le rendement des sols et répondre aux attentes des consommateurs. Dans sa croissance, la pomme de terre a besoin de grandes quantités de nutriments, en particulier N (azote), P (phosphore) et K (potassium) [3], ce qui justifie l'utilisation de ces substances par les agriculteurs pour fertiliser leurs sols. Ces engrais améliorent le produit agricole au niveau quantitatif mais malheureusement au niveau qualitatif, les pommes de terre subissent des modifications de leurs propriétés physiologiques végétatives [4]–[7]. Plusieurs études menées sur les pommes de terre ont confirmé une grande différence entre les pommes de terre biologiques (traitées traditionnellement) et les pommes de terre produites à partir de champs traités avec les engrais NPK dans leur contenu nutritionnel. Les pommes de terre biologiques qui sont plus saines ont

une teneur élevée en protéines et en acides aminés, en vitamine C et en matière sèche, tout en étant faibles en azote, en nitrates et en sucres réducteurs. Sans oublier l'impact de l'utilisation excessive des engrais sur l'environnement, que ce soit sur leur niveau de production qui conduit à la pollution de l'air ou sur leur niveau d'utilisation qui conduit à la pollution des nappes phréatiques [8].

Cette étude vise à distinguer entre les pommes de terre produites à partir d'un champ traité traditionnellement avec le fumier des moutons et des ânes et celles produites à partir du champ traité avec de l'engrais NPK.

I.1. Préparation des échantillons

Dans ces expériences, des tubercules de pomme de terre du genre *Solanum tuberosum* L. ont été utilisés en raison de leur large consommation dans le monde entier. Les pommes de terre ont été prélevées sur deux types de sol, des échantillons d'un champ traité traditionnellement avec le fumier de moutons et d'ânes et des échantillons d'un champ traité avec des engrais NPK. Pour le champ de pomme de terre traitée avec les engrais NPK, 4 applications de fertilisation ont été faites : 30 jours après la plantation, la première application était de 0,2 tonne de 20-20-20 par hectare. Les applications deuxième, troisième et quatrième après successivement 50, 70 et 90 jours de plantation étaient de 0,5 tonne chacune avec un taux de 14-7-21+2MgO par hectare. Et pour le champ traité traditionnellement environ de 2 tonnes par hectare du fumier de moutons et d'ânes a été utilisé.

Les échantillons ont été soigneusement choisis, ils ont été d'abord, inspectés visuellement et les échantillons qui présentaient des signes d'infection ou des taches vertes de chlorophylle ont été discriminés. Ensuite, elles ont été préalablement lavées et séchées avant d'être coupées en frites. Des échantillons de 100 g ont été placés dans des erlenmeyers en verre scellés (avec un bouchon en caoutchouc) étiquetés de 1 litre (figure 5-1) et conservés dans un endroit à une température ambiante de 25 ± 2 °C pendant une semaine (7 jours) avant d'être utilisé dans les expériences.



Figure 5-1 Exemples des échantillons préparés et utilisés

I.2. Réponses temporelles des capteurs

Pour cette application cinq capteurs ont présenté une grande sensibilité à la présence des échantillons des pommes de terre sont : MQ-7, MQ-136, TGS-821, TGS-822 et TGS-2620.

Deux séries d'expériences ont été réalisées, vingt expériences par type d'échantillons pour construire des informations solides. Une première série concernant les pommes de terre récoltées à partir du champ traité traditionnellement et une deuxième série concernant celles récoltées du champ traité avec des engrais NPK. Les données pour chaque type d'échantillon ont été prélevées pendant 8 minutes avec un démarrage du processus après 30 secondes pour vérifier la stabilité des capteurs tout en laissant à la matrice des capteurs un temps de repos de 10 minutes pour permettre aux capteurs de revenir à leurs états initiaux à la fin de chaque expérience.

Les courbes de réponse des capteurs sont illustrées à la figure 5-2 où nous avons choisi de projeter seulement cinq échantillons de chaque type de pomme de terre. L'axe horizontal indique le temps en secondes, l'axe vertical indique la valeur de la tension relative de sortie des capteurs en volts.

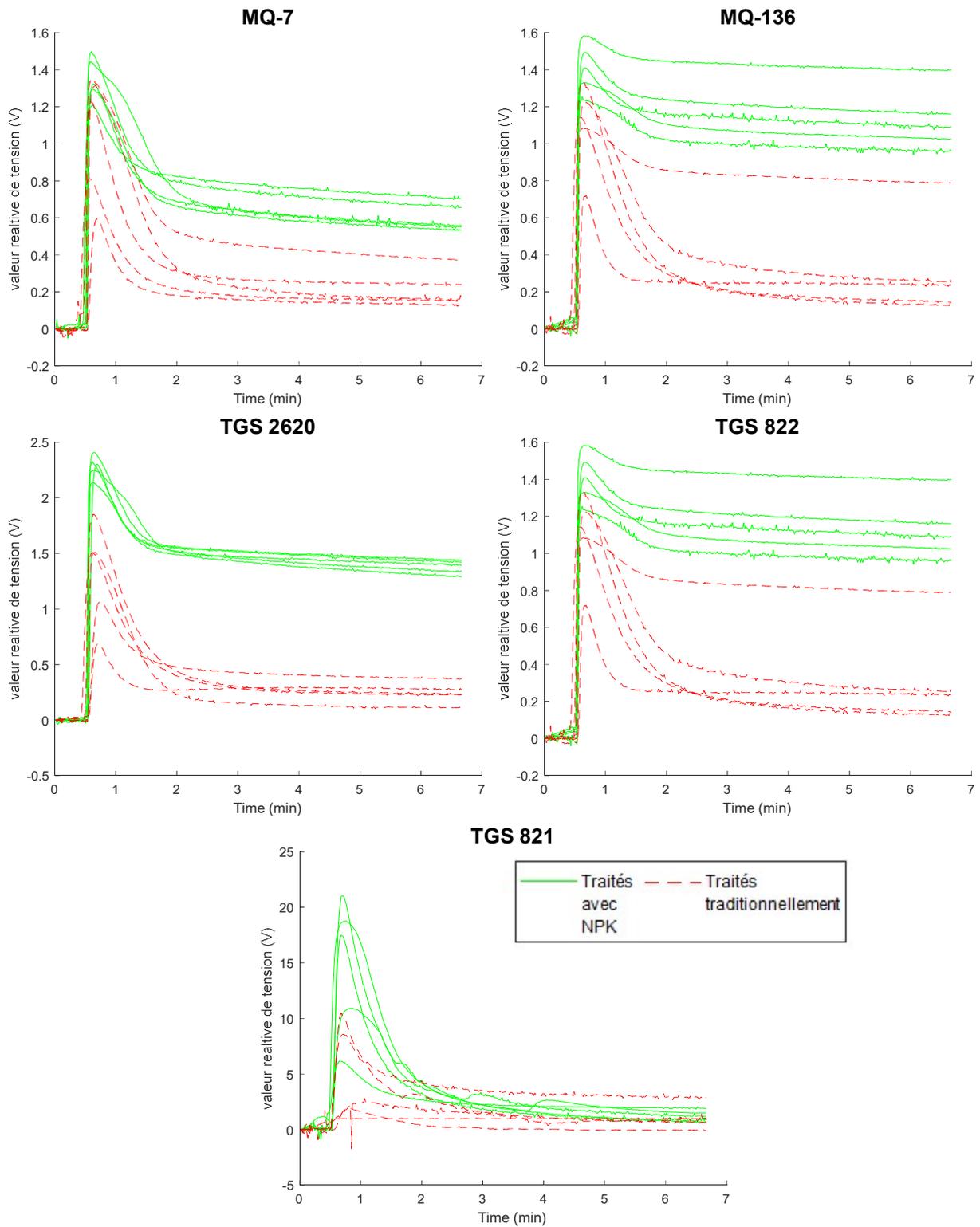


Figure 5-2 Exemples des réponses enregistrées par les cinq capteurs à 5 échantillons de pommes de terre prélevés sur le champ traité traditionnellement et à 5 autres échantillons prélevés sur le champ traité par les engrais NPK

À partir de la figure 5-2, nous remarquons, dans un premier temps, que la tension augmente lorsque nous faisons passer le flux d'air enrichi par l'odeur des pommes de terre dans la chambre du capteur. Presque tous les capteurs ont le même comportement pour chaque mesure ce qui peut s'expliquer par la détection des cinq capteurs de la présence des échantillons de pommes de terre. Ce qui est remarquable, c'est la façon dont les capteurs répondent à chaque type d'échantillons, les capteurs répondent avec une tension de sortie très élevée aux échantillons de pommes de terre récoltées d'un champ traité avec les engrais NPK, contrairement à ceux récoltés d'un champ traité traditionnellement. Cette différence indique que le système multisensoriel peut distinguer les échantillons en fonction de leurs réponses.

I.3. Distinction des pommes de terre selon leurs champs de culture

Après la même étape de prétraitement des données adoptée dans le chapitre précédent, notre matrice de données sera composée de 15 colonnes relatives à 3 caractéristiques (la valeur maximale, l'aire et la valeur stabilisée) par chacun des cinq capteurs et 40 lignes relatives à 40 expériences, 20 expériences par type d'échantillon.

I.3.1. Résultats de l'ACP

Cette méthode est utilisée pour réduire notre espace de données de quinze dimensions à un espace à deux dimensions (PC1 et PC2). L'objectif est de visualiser les données dans un graphe sans perdre beaucoup d'informations et de se faire une idée sur une éventuelle distinction entre les échantillons. Le résultat de la méthode PCA est illustré à la figure 5-3.

Le nuage de points en deux dimensions montre que les nuages de points de chaque groupe des échantillons de pommes de terre sont bien corrélés et regroupés et se chevauchent légèrement, ce qui permet d'identifier visuellement les groupes des individus correspondants aux échantillons relatifs au champ traité traditionnellement et ceux relatif au champ traité avec les engrais NPK. Le taux de contribution cumulé des deux premières composantes principales était d'environ 98%, ce qui indique que le plan à deux dimensions PC1 et PC2 est assez représentatif et une perte minimale d'informations quasiment négligeable. Le premier axe PC1 qui représente 94,05% de la variation est celui qui distingue le plus la nature de l'échantillon de la pomme de terre.

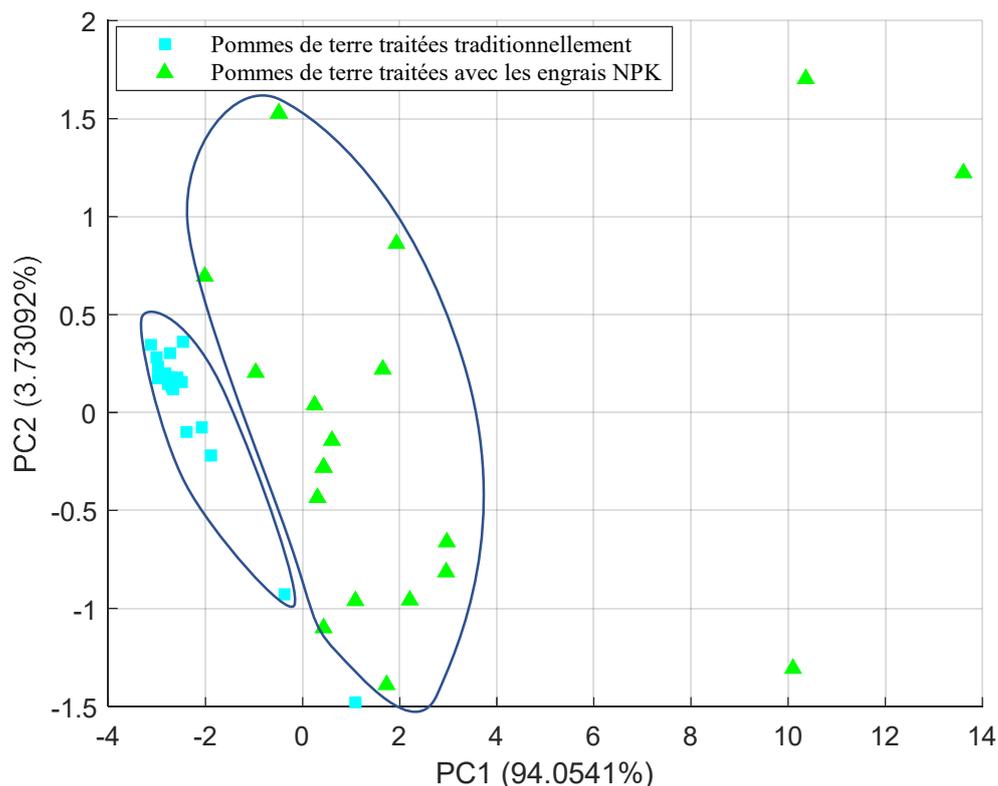


Figure 5-3 Nuage de points PCA des échantillons de pommes de terre selon leurs sources, traités traditionnellement ou traités avec des engrais NPK

Ce bon résultat prouve qu'il est possible de distinguer visuellement les échantillons de pomme de terre selon leur provenance (traitée traditionnellement avec le fumier de moutons et d'ânes ou traitée avec des engrais NPK). Comme évoqué précédemment, la méthode ACP se limite dans la génération des représentations graphiques des données, son principal inconvénient est l'absence d'une règle de décision qui permet l'identification des échantillons inconnus. Pour ce faire nous allons exploiter la méthode SVM par la suite.

I.3.2. Résultats du SVM

Etant une méthode supervisée, la méthode SVM nécessite la définition des classes d'appartenance à priori. Pour notre cas, les classes sont connues à l'avance car c'est nous qui préparons les échantillons pour les expériences. L'approche SVM se compose de deux phases, comme décrit au chapitre II, la première phase est l'apprentissage, il s'agit de la construction du modèle de classification, et la deuxième phase est la validation ou le test. En conséquence, les données contenues dans l'ensemble de données doivent être divisées en deux groupes : un pour

l'apprentissage et l'autre pour la validation. Dans la même optique, nous avons opté pour une validation croisée quintuple (5-fold), 80% pour l'apprentissage et 20% pour le test.

Le résultat de la méthode SVM est montré sur la figure 5-4 dans laquelle nous avons choisi les moyennes stabilisées du capteur MQ7 et du capteur TGS821 pour visualiser les résultats dans un plan bidimensionnel.

En observant le résultat de l'identification, un très bon taux d'identification de 92,5% a été atteint, reflété dans l'affectation correcte des échantillons à leurs classes de données, à l'exception de trois erreurs de reclassement, deux mesures du groupe traité traditionnellement et une mesure du groupe traité avec les engrais NPK.

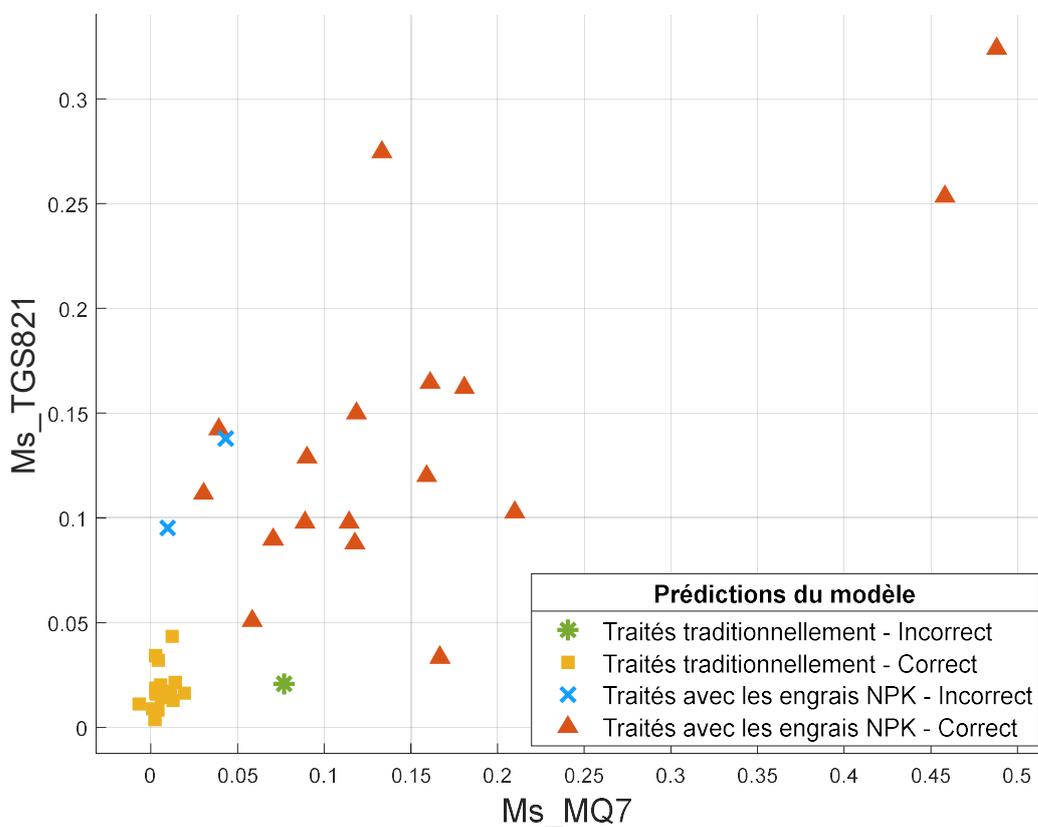


Figure 5-4 Résultat de la classification SVM pour les échantillons de pommes de terre selon leurs sources, non traitées ou traitées avec des engrais NPK

Ce résultat prometteur, tel que celui que nous avons réalisé dans nos premières études sur la menthe à l'aide du même dispositif [9]–[11], renforce l'utilité de notre approche proposée qui est simple et peu coûteuse, qu'il s'agisse de distinguer les menthes traitées aux insecticides de celles non traitées ou pour prédire les sources de pommes de terre provenant d'un champ

traité traditionnellement avec le fumier des moutons et des ânes ou d'un champ traité par des engrais NPK.

I.4. Conclusion

Dans ce chapitre, nous avons étudié la discrimination entre deux types de pommes de terre selon leurs sources : champ traité avec les engrais NPK et champ traité traditionnellement avec le fumier de moutons et d'ânes, ce légume largement consommé partout dans le monde est aussi soumis à des traitements par des substances chimiques, ce qui rend sa consommation dangereuse pour la santé humaine. Pour atteindre cet objectif, un nez électronique composé d'un système multi-capteurs couplé à des algorithmes d'apprentissage automatique a été appliqué comme méthode de détection alternative pour évaluer l'origine des pommes de terre en fonction de la nature du champ. Des échantillons de pommes de terre ont été prélevés d'un champ traité traditionnellement avec le fumier des moutons et des ânes et d'autres échantillons d'un champ traité avec des engrais NPK. La méthode de la machine à vecteurs de support (SVM) a été utilisée pour classer les données et elle a obtenu des meilleures performances avec une précision de 92,5%. Enfin, le système s'est révélé efficace pour distinguer entre les deux types de pommes de terre analysées et il est possible qu'il soit exploité comme une technique alternative pour vérifier les pommes de terre biologique (traitées traditionnellement).

II. Deuxième Partie : DISCRIMINATION DES POMMES DE TERRE A L'AIDE D'UN SYSTEME MULTISENSORIEL ET D'UNE METHODE MICROBIOLOGIQUE

Comme on l'a évoqué dans la première partie du présent chapitre, l'augmentation du rendement des terres agricoles a toujours été une grande préoccupation pour les agriculteurs afin d'augmenter leurs profits, en particulier avec l'explosion démographique et la forte demande de produits agricoles, de sorte que la plupart des agriculteurs ont recours à des engrais organiques et chimiques. L'ajout de ces engrais conduit à des rendements accrus et à une amélioration de la quantité, mais aussi inévitablement à un changement significatif dans les composants du produit [12]–[16]. Quant aux engrais chimiques, ils ont un impact important et des répercussions directes sur les composants vivants, dont la santé des humains [17] et les plantes [18]. Il affecte également négativement les composants de l'écosystème, l'eau, l'air et le sol [19]. Quant à l'utilisation d'engrais d'origine animale, autrefois considérés comme organiques, aujourd'hui la situation a changé car il a été prouvé qu'ils contiennent des métaux lourds [20]. De plus, ces engrais ont subi une modification importante de leur teneur en raison de l'utilisation excessive de médicaments dans l'engraissement et le traitement des animaux.

Des études ont prouvé que les produits agricoles contiennent ces substances dangereuses pour la santé humaine [21].

Pour distinguer les produits agricoles conventionnels des produits agricoles biologiques, des études et des analyses ont été réalisées à l'aide d'équipements physico-chimiques complexes coûteux et lourds [22]–[26] ou de techniques biologiques [27], [28] nécessitant un temps de manipulation important. Pour résoudre le problème du manque d'un outil simple, portable et moins cher, nous avons eu recours à un nez électronique à base de capteurs de gaz à oxyde métallique pour discriminer les pommes de terre selon la nature des champs cultivés (champ non traité, traité avec des engrais NPK ou traité avec du fumier de poulet).

II.1. Préparation des échantillons

Dans cette présente étude, nous avons adopté la même démarche dans la préparation des échantillons de tests, sauf que cette fois nous avons utilisé trois types d'échantillons de pomme de terre, des tubercules d'un champ traité avec des engrais traditionnels (fumier de moutons et d'ânes), d'un champ traité avec des engrais NPK et d'un champ traité avec du fumier de poulet.

Pour le champ de pomme de terre traité avec des engrais NPK, la fertilisation a été faite comme suit : la première application était de 0,2 tonne de 20-20-20 par hectare trente jours après la plantation. Les deuxième, troisième et quatrième application étaient de 0,5 tonne chacune avec un taux de 14-7-21+2MgO par hectare après successivement 50, 70 et 90 jours. Pour le champ de pommes de terre traité avec du fumier de poulet, la quantité était de 2 tonnes par hectare de même pour le champ traité avec les engrais traditionnels (fumier de moutons et d'ânes).

II.2. Analyse microbiologique

II.2.1. Principe de la méthode de dénombrement de la population microbiologique totale

Cette étude a été réalisé en collaboration avec Madame Asmae Aboukacem Professeure à l'Ecole Supérieure de Technologie (EST) de Khenifra. Des échantillons de 25 g de pommes de terre récoltées de chaque type de champ (traitées avec les engrais traditionnels, traitées avec du fumier de poulet ou avec des engrais NPK) ont été prélevés et placés dans un sac Stomacher stérile contenant 225 ml d'eau peptonnée Tamponnée (BPW, Biokar Diagnostics, France). Ensuite et pendant 2 minutes, l'échantillon et l'eau peptonnée tamponnée ont été mélangé, ainsi que la préparation des dilutions décimales en utilisant le même diluant. Ces dilutions ont ensuite

été incorporées dans une masse de 18 ml d'un milieu Plate Count Agar (PCA, Biokar Diagnostics, France) dans des boîtes de Pétri de 90 mm de diamètre. Ces boîtes ont été incubées pendant 3 jours à 30°C comme recommandé par la norme ISO 4833-1. Le comptage des colonies a été réalisé conformément à la norme ISO 7218. Le nombre de colonies présentes est calculé en log₁₀ unités formatrices de colonies par gramme d'échantillon (CFU/g), c'est le nombre total de colonies viables (TVC) qui est une estimation quantitative de la concentration de micro-organismes.

II.2.2. Résultats et discussion

Les résultats des analyses microbiologiques réalisées sur les pommes de terre sont illustrés sur la figure 5-5. Cette figure montre l'évolution de microorganismes totaux en log₁₀ CFU/g dans trois catégories de pommes de terre en fonction du nombre de jours de stockage à température ambiante.

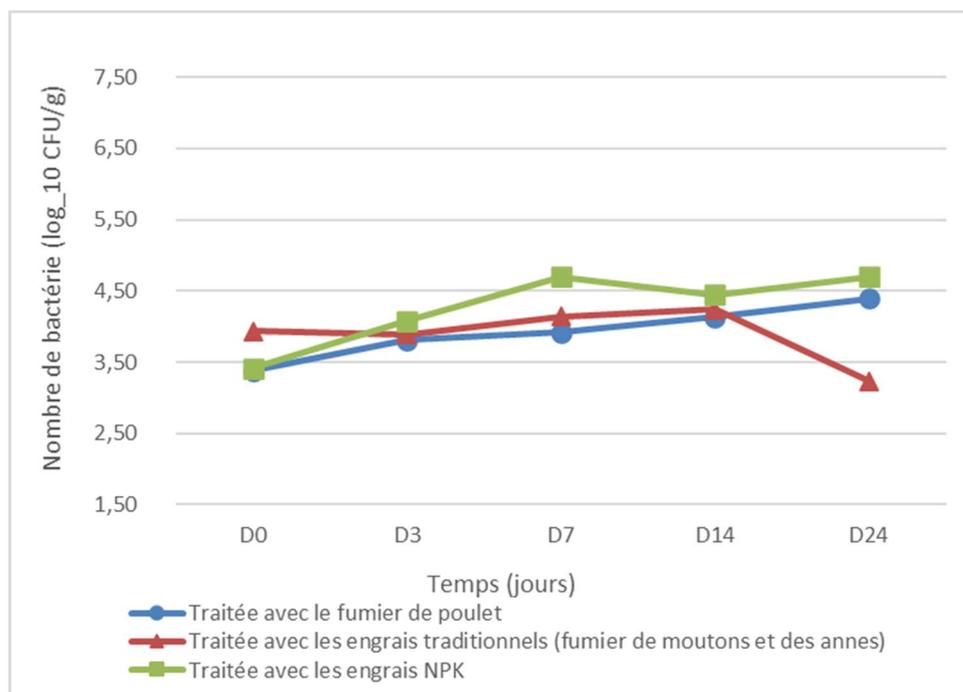


Figure 5-5 Analyses microbiologiques des échantillons de pommes de terre

Le but de cette étude est de comparer le comptage des unités formant colonie (UFC) des trois types de pommes de terre selon leurs provenances (récoltées d'un champ traité avec des engrais traditionnels, traité avec du fumier de poulet ou traité avec des engrais NPK), et d'explorer si ces tests peuvent détecter une sous-population de pommes de terre différenciée selon le type de traitement qu'elles ont subi.

Le développement des microorganismes a montré une tendance comparable pour les pommes de terre récoltées du champ traité avec du fumier de poulet et celles récoltées du champ traité avec des engrais NPK en enregistrant une supériorité par cette dernière catégorie, en particulier après 7 jours de stockage. Le nombre de micro-organismes n'a cessé d'augmenter depuis le début de l'expérimentation jusqu'à la fin (jour 24) cette valeur maximale était d'environ 4,5 UFC/g pour ces deux types de pomme de terre. En revanche, les pommes de terre récoltées du champ traité avec des engrais traditionnels présentaient un nombre d'unités formant des colonies plus élevé au départ qui a connu une légère augmentation du jour 3 au jour 14 pour rejoindre les autres catégories avant de chuter remarquablement à partir du 14^e jour de stockage. Ce nombre a connu une baisse continue jusqu'à la fin de l'expérience (jour 24) avec une valeur minimale d'environ 3,1 UFC/g.

En étudiant les courbes, on peut dire que jusqu'au 14^{ème} jour il est difficile de distinguer les échantillons car les signaux sont proches et se chevauchent, notamment la courbe des pommes de terre traitées avec des engrais traditionnels et celle des pommes de terre traitées au fumier de poulet qui sont si proches qu'il est difficile de les distinguer, contrairement à la courbe des pommes de terre traitées aux engrais NPK qui est relativement éloignée. Mais après le quatorzième jour, là, nous remarquons qu'il est possible de distinguer les pommes de terre traitées avec des engrais traditionnels des autres, cependant, la nature du traitement ne peut pas être jugée car les deux courbes des pommes de terre traitées avec le fumier de poulet et traitées avec les engrais NPK sont très proches et ont presque la même réponse.

II.3. Analyse avec le nez électronique

II.3.1. Préparation des échantillons

La méthodologie adoptée pour préparer des échantillons de tubercules de pomme de terre du genre *Solanum tuberosum* L. est la suivante : trois types de pommes de terre ont été prélevés dans trois champs distincts : le premier considéré biologique traité avec des engrais traditionnels (fumier de moutons et d'ânes), le second traité avec des engrais NPK et le troisième traité avec du fumier de poulet. Suite à une inspection visuelle des échantillons, les tubercules endommagés lors de la récolte ou présentant des signes d'infection ou qui n'étaient pas encore mûrs et qui présentaient des taches vertes de chlorophylle ont été retirés.

Après lavage et séchage des tubercules, celles-ci ont été coupées en frites avant de placer 100g de chaque échantillon dans des erlenmeyers en verre scellés (avec un bouchon en

caoutchouc) de volume 1 litre portant des étiquettes qui renseignent sur la nature de l'échantillon.

Les échantillons préparés ont été placés dans une pièce à une température ambiante de $25 \pm 2^\circ\text{C}$ pendant 7 jours avant d'être analysés. La figure 5-6 montre une photo de quelques échantillons utilisés.



Figure 5-6 Exemples des échantillons préparés et utilisés

II.3.2. Résultat de l'analyse multisensorielle

II.3.2.1 Réponses du réseau de capteurs et prétraitement

Après avoir mis les différents échantillons de pommes de terre sélectionnées dans leurs enclos sellés pendant une semaine, ils ont été soumis au processus de mesure. Ce processus a été répété 20 fois pour chaque type de pomme de terre, pour un total de 60 tests afin de vérifier la répétabilité. Pendant le processus de mesure, les espaces de tête des échantillons poussés par le ventilateur à la chambre des capteurs ont engendré les courbes de réponses illustrées sur la figure 5-7.

La figure 5-7 représente les courbes de réponse en tension relative en fonction du temps de cinq échantillons choisis aléatoirement de chaque type de pomme de terre.

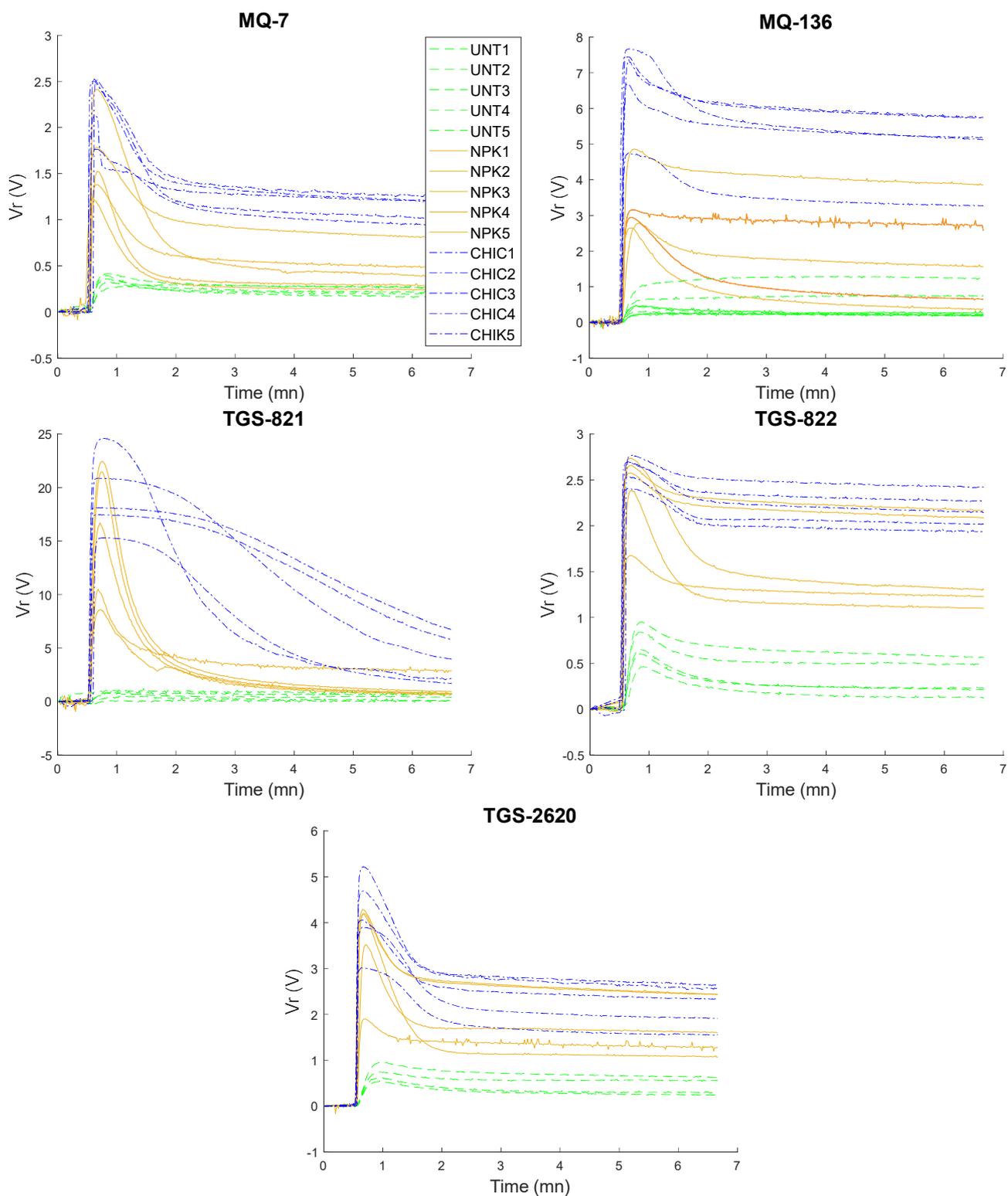


Figure 5-7 Exemples de réponses enregistrées par les cinq capteurs à cinq échantillons de pommes de terre de chaque type, la couleur verte représente celles prélevées du champ traité avec des engrais traditionnels, la couleur orange pour celles traitées avec des engrais NPK et la couleur bleue pour celles traitées avec du fumier de poulet

La première remarque tirée de la figure est que les cinq capteurs répondent positivement à la présence des échantillons de pomme de terre avec des degrés différents. La tension de réponse des capteurs augmente une fois que le flux d'air enrichi par l'espace libre des échantillons de pomme de terre est injecté dans la chambre des capteurs. Ce qui est très remarquable c'est la valeur avec laquelle tous les capteurs répondent, tous les capteurs sans exception répondent avec une très haute tension aux échantillons traités traditionnellement par rapport à traités différemment. On peut alors compter sur la valeur maximale de la tension relative pour distinguer les échantillons traités avec des engrais traditionnels des autres, ce qui ne nous aidera pas à distinguer les échantillons traités avec des engrais NPK de ceux traités avec du fumier de poulet sauf peut-être pour le cas du capteur MQ-136. A cet effet, et après examen des graphes, il a été remarqué que les réponses de ces deux échantillons (traités avec des engrais NPK et traités avec du fumier de poulet) se distinguaient par l'air occupé par leurs signaux et les valeurs stabilisées à la fin de chaque expérience.

Les différences enregistrées par le système multisensoriel peuvent être utilisées pour distinguer les échantillons de pommes de terre en fonction de leurs réponses. Ces différences seront utilisées comme des caractéristiques des courbes de réponse originales à utiliser dans la reconnaissance de formes afin d'obtenir un résultat de reconnaissance optimal. Dans notre cas, il existe trois caractéristiques, à savoir la valeur maximale, la surface occupée par le signal, et la valeur stabilisée du signal entre les 450s et 500s comme choisi dans toutes les expériences précédentes.

Cinq ensembles de données ont été créés :

- X1 : un ensemble de données de pommes de terre récoltées du champ traité traditionnellement avec des engrais de moutons et d'ânes.
- X2 : un ensemble de données de pommes de terre récoltées du champ traité avec des engrais NPK.
- X3 : un ensemble de données de pommes de terre récoltées du champ traité avec du fumier de poulet.
- X4 : un ensemble de données des pommes de terre traitées avec le fumier de poulet et des engrais NPK à la fois (X2+X3).
- X5 : toutes les données.

La matrice résultante est donc composée de quinze colonnes : 5 réponses de capteurs multipliées par 3 caractéristiques principales pour chacune, et de soixante lignes : 20 échantillons multipliés par 3 types de pommes de terre.

II.3.2.2 Résultat de la classification

La classification a été effectuée à l'aide de l'algorithme SVM. Comme déjà évoqué, SVM est une méthode efficace pour classer les ensembles de données non linéaires et de petits. En exploitant cette méthode, nous sommes à la recherche du développement d'une règle de décision sûre pour la discrimination et l'identification des échantillons de pommes de terre selon la nature du champ d'où ils sont récoltés : traité avec des engrais traditionnels, traité avec les engrais NPK ou avec le fumier de poulet.

Etant donné que cette méthode choisie est supervisée, le choix des groupes de données d'apprentissage se fait à priori. Dans ce cas, et comme les cas précédents, nous nous appuyons sur les cibles préalablement connues avant l'analyse.

L'utilisation des machines à vecteurs de support (SVM) comporte deux phases. La création du modèle de classification est la phase initiale. Tandis que la deuxième phase est la validation du modèle construit. Ainsi, les données doivent être divisées en deux sous-ensembles, que nous surnomons respectivement l'ensemble de données d'apprentissage et l'ensemble de données de validation. Pour la validation du modèle, une validation croisée 5-fold a été choisie puisque cette méthode de validation protège contre le surentraînement en divisant l'ensemble de données en cinq plis de même taille et en estimant la précision sur chaque pli. Par conséquent, les données ont été à chaque fois divisées en 80 % pour l'entraînement et 20 % pour les tests en assurant que toutes les données passeront par l'entraînement et le test. Pour la fonction noyau, toutes les fonctions ont été testées et il s'avère que la mieux adoptée pour ce cas est la fonction linéaire.

Dans les figures 5-8, 5-9 et 5-10 et pour la visualisation, la valeur maximale extraite du capteur MQ-7 et l'aire extraite du capteur TGS821 ont été choisis afin de visualiser les résultats dans un plan bidimensionnel.

La classification a été faite selon deux manières : la première manière qu'on peut aussi appeler une méthode directe où toutes les données (X5) ont été injectées en une seule fois, la figure 5-8 montre les résultats obtenus, la figure 5-8 (a) représente la projection des données et

(b) la matrice de confusion. Cette méthode a obtenu un taux de réussite d'environ 92 % (exactement 91,7 %). Seuls 5 des 60 échantillons ont été mal classés, deux appartenant au premier groupe concernant les pommes de terre récoltées dans des champs traités avec des engrais traditionnels, un appartenant au deuxième groupe concernant les pommes de terre

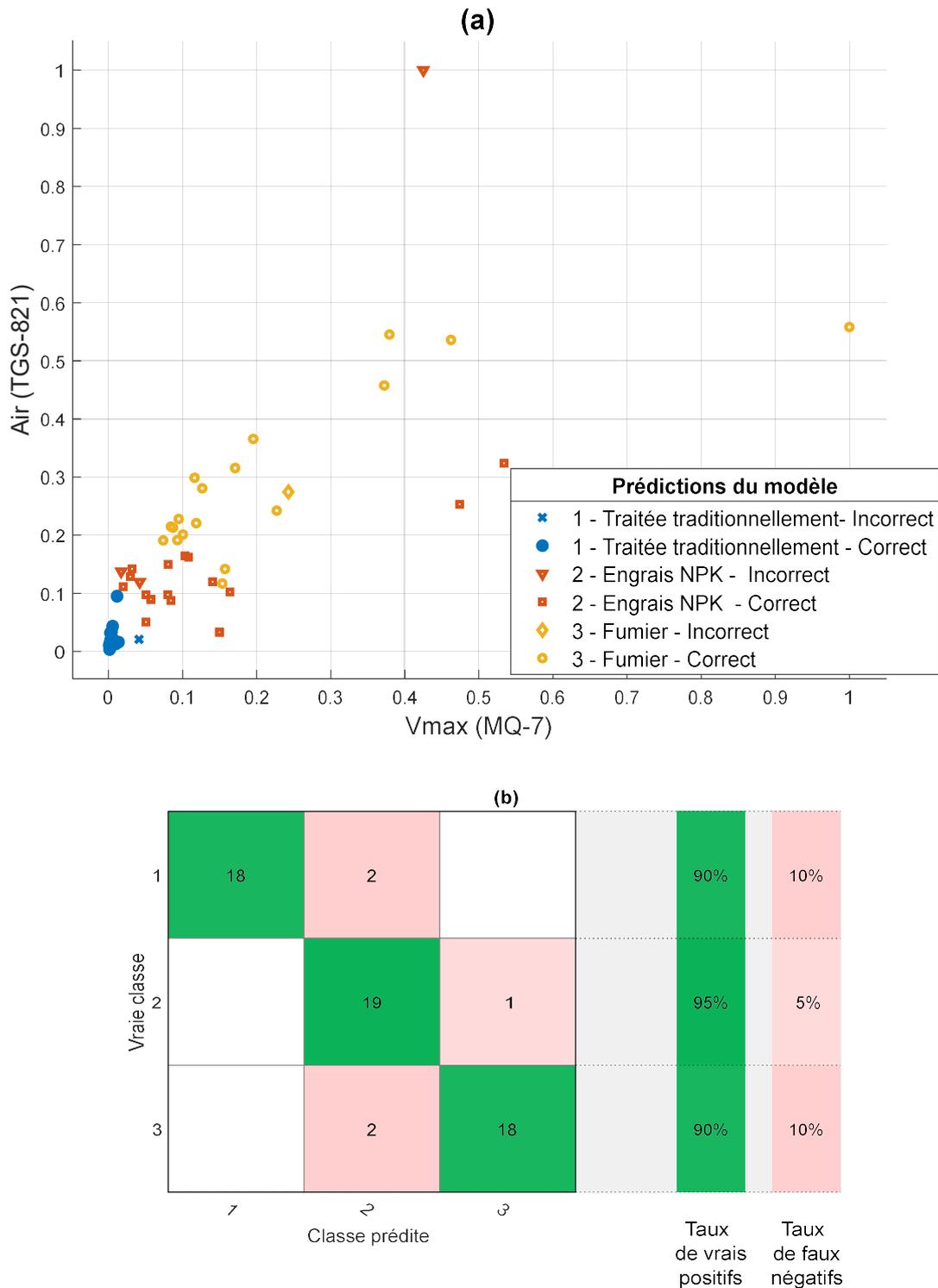


Figure 5-8 Résultats SVM pour l'identification des types de pommes de terre par la première approche, (a) représentation graphique du résultat et (b) matrice de confusion

récoltées dans des champs traités avec des engrais NPK, et les deux autres sont liés au troisième et dernier groupe qui représente les échantillons de pommes de terre traitées avec du fumier de poulet.

Pour améliorer encore les résultats, une autre approche peut être utilisée, cette approche (une méthode composée) se compose de deux étapes, la première consiste à identifier les échantillons traités avec le fumier de poulet et les engrais NPK (X4) de ceux traités traditionnellement (X1), et la deuxième étape est l'identification de la nature du traitement s'il s'agit d'engrais NPK (X2) ou de fumier de poulet (X3). Les figures 5-9 et 5-10 illustrent les résultats de cette seconde voie, (a) et (c) sont successivement les projections de données de pommes de terre traitées traditionnellement contre les autres traitées différemment et les données relatives aux pommes de terre traitées avec les engrais NPK contre celles traitées avec le fumier de poulet. (b) et (d) sont aussi les matrices de confusion de chaque cas.

Pour l'identification des échantillons de pommes de terre traitées traditionnellement de celles traitées différemment, l'algorithme SVM a obtenu un taux de succès de 96,7 % avec seulement deux échantillons mal classés, un de chaque groupe.

Pour la reconnaissance de la nature du traitement, un taux de réussite de 92,5% a été atteint, avec trois échantillons mal classés, deux appartenant au groupe des traités avec les engrais NPK et un au groupe des traités avec le fumier de poulet.

Pour résumer les résultats de la classification, les deux méthodes sont bonnes à utiliser sauf que la méthode directe fait gagner du temps mais avec un taux de succès relativement faible (91,7%), et la deuxième méthode a réalisé des bons taux de (96,5% et 92,5%) toutefois elle demande un peu plus de temps de traitement.

Ce résultat atteint permet de dire qu'il est possible d'utiliser un système multi-capteurs simple et peu coûteux pour discriminer les pommes de terre selon la nature des champs s'ils sont traités traditionnellement ou traités avec des engrais NPK ou du fumier de poulet.

Enfin, contrairement à de nombreuses études qui ont examiné la relation entre les données microbiologiques et celles du nez électronique [29], [30]. Notre étude présente une enquête qui a été menée sur les deux méthodes et il a été constaté que si nous comptons sur l'analyse microbiologique pour déterminer la nature de la pomme de terre, cela prendra beaucoup de temps et d'efforts (3 jours de préparation des échantillons pour le test et 24 jours

de surveillance et d'observation). Par opposition à l'utilisation d'un système multi-capteurs qui se caractérise par son faible coût, sa simplicité d'utilisation et sa capacité à économiser du temps et des efforts.

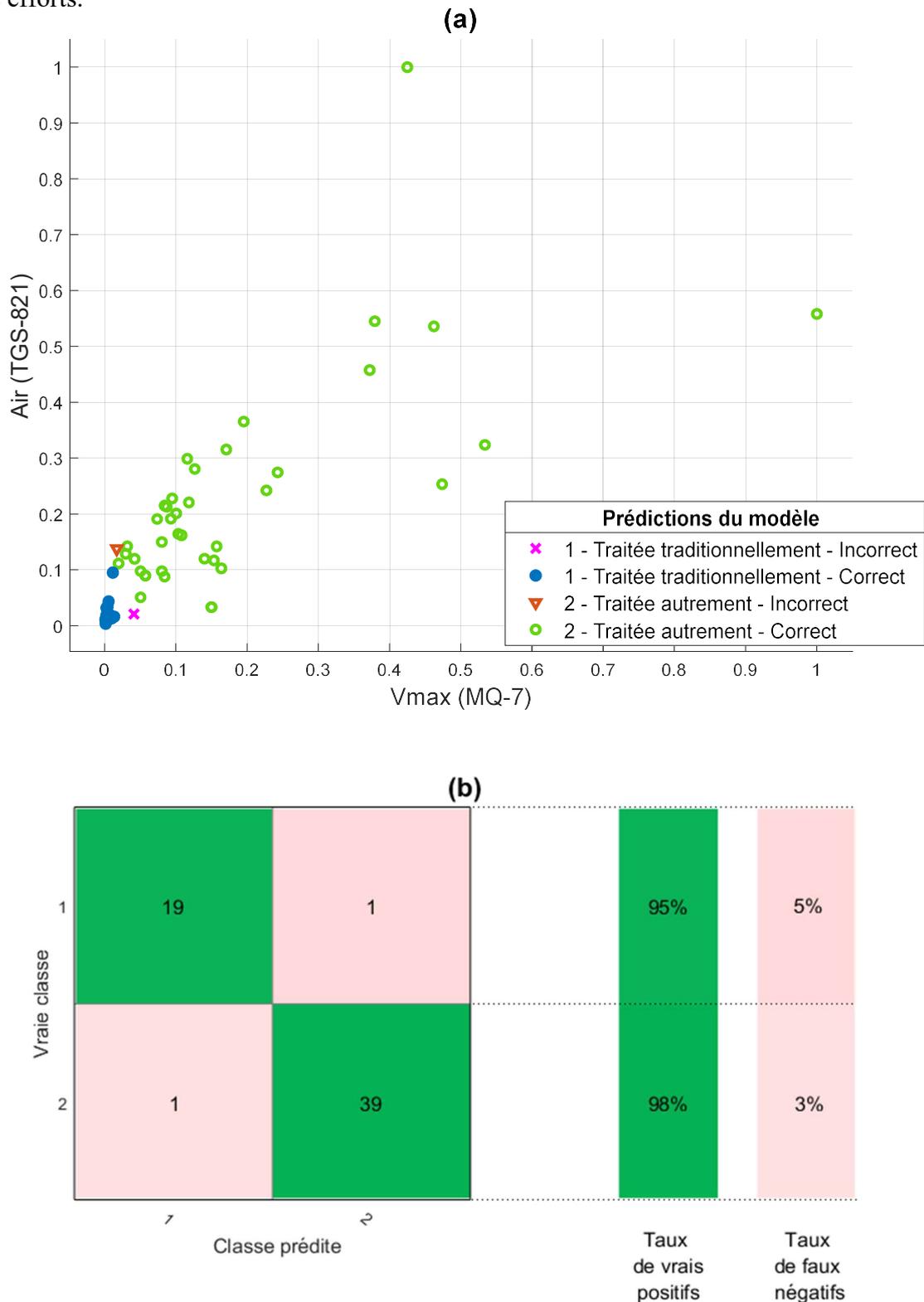


Figure 5-9 Résultat de la première étape de la deuxième approche, (a) représentation du résultat de la discrimination des pommes de terre traitées traditionnellement de celles traitées différemment et (b) la matrice de confusion

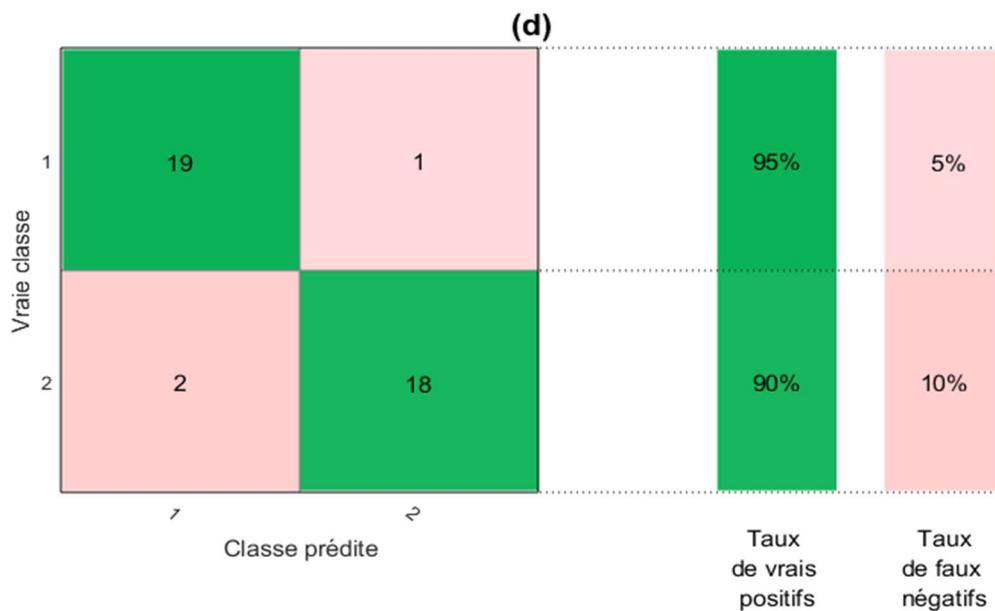
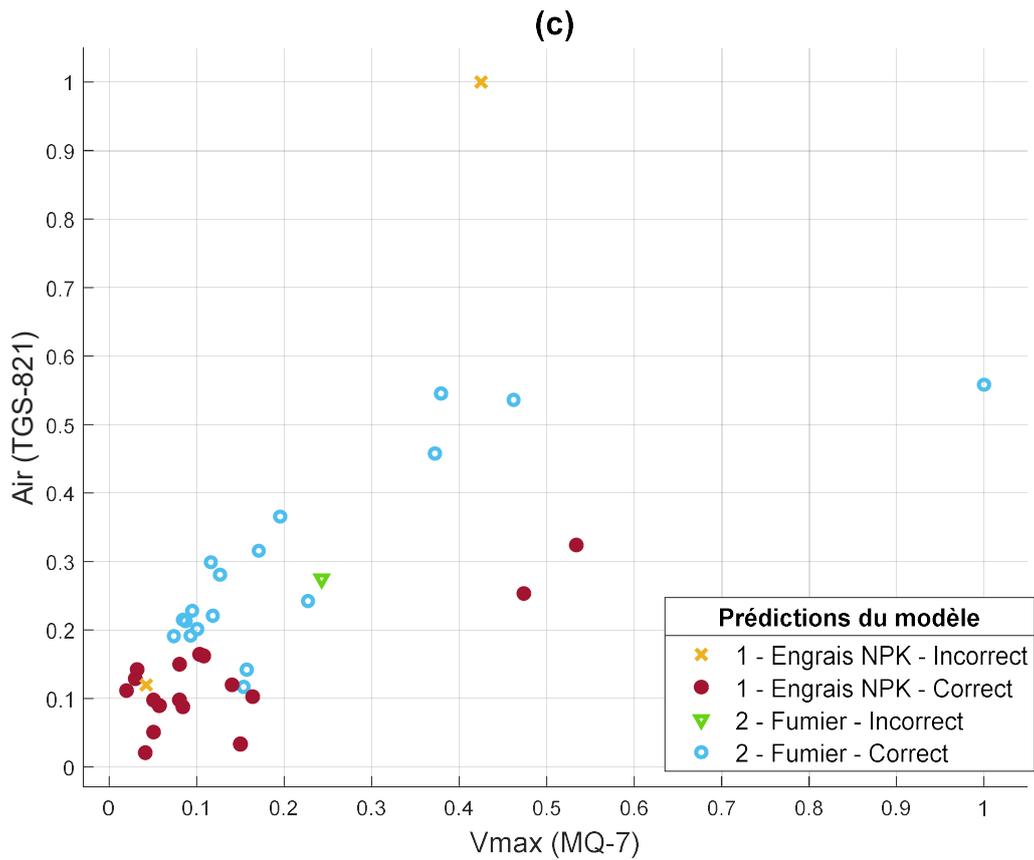


Figure 5-10 Résultat de la deuxième étape de la deuxième approche, (c) représentation du résultat de la discrimination des pommes de terre traitées avec du fumier de poulet et de celles traitées avec des engrais NPK et (b) la matrice de confusion

II.4. Conclusion

L'analyse microbiologique a été utilisée comme méthode complémentaire pour mettre en évidence la différence entre les échantillons de pommes de terre selon la nature des champs dans lesquels elles ont été récoltées, et une grande différence a été enregistrée entre les échantillons en ce qui concerne le nombre de colonies de bactéries, surtout après le quatorzième jour du test, avec une petite différence enregistrée le septième jour qui a été utilisée pour ouvrir la voie à l'analyse du nez électronique.

Après avoir préparé les échantillons de pommes de terre de différents types selon la nature des traitements du champ, leurs informations ont été collectées à l'aide d'un système multi-capteurs, par la suite, la méthode d'apprentissage automatique des machines à vecteurs de support (SVM) a été utilisée et elle a prouvé son efficacité à distinguer les sources de pommes de terre dans deux manières différentes : une méthode dite directe où toutes les données ont été utilisées à la fois, un taux de réussite de 92% a été atteint. Une méthode composée comportant deux étapes : la première consiste à identifier les échantillons traités traditionnellement de ceux traités autrement, et la deuxième étape, l'identification du traitement si c'est avec des engrais NPK ou du fumier de poulet, les résultats étaient de 96,7% pour la première étape et 92,5 % pour la seconde.

Sur la base de ces résultats, notre dispositif facile à manipuler et portable peut être utilisé comme un outil efficace pour déterminer la nature des pommes de terre en fonction de la nature du champ dans lequel elles ont été récoltées.

Références

- [1] H. Zhang, F. Xu, Y. Wu, H. hai Hu, and X. feng Dai, "Progress of potato staple food research and industry development in China," *J. Integr. Agric.*, vol. 16, no. 12, pp. 2924– 2932, 2017, doi: 10.1016/S2095-3119(17)61736-2.
- [2] K. A. Beals, "Potatoes , Nutrition and Health," 2018.
- [3] P. Brinkman et al., "Identification and prospective stability of electronic nose (eNose)–derived inflammatory phenotypes in patients with severe asthma," *J. Allergy Clin. Immunol.*, vol. 143, no. 5, pp. 1811-1820.e7, 2019, doi: 10.1016/j.jaci.2018.10.058.
- [4] M. M. Abd El-Azeim, M. A. Sherif, M. S. Hussien, I. A. A. Tantawy, and S. O. Bashandy, "Impacts of nano- And non-nanofertilizers on potato quality and productivity," *Shengtai Xuebao/ Acta Ecol. Sin.*, vol. 40, no. 5, pp. 388–397, 2020, doi: 10.1016/j.chnaes.2019.12.007.

- [5] M.M.E. Ali, S.A. Petropoulos, D.A.F.H. Selim, M. Elbagory, M.M. Othman, A.E.-D. Omara, M.H. Mohamed, “Plant growth, yield and quality of potato crop in relation to potassium fertilization,” *Agronomy*, vol. 11, no. 4, pp. 1–16, 2021, doi: 10.3390/AGRONOMY11040675.
- [6] A. S. Petropoulos, A. Fernandes, N. Polyzos, V. Antoniadis, L. Barros, and C. F. R. I. Ferreira, “The Impact of Fertilization Regime on the Crop Performance and Chemical Composition of Potato,” *Agronomy*, vol. 10, pp. 1–18, 2020.
- [7] K. Mokrani, K. Hamdi, and N. Tarchoun, “Potato (*Solanum Tuberosum* L.) Response to Nitrogen, Phosphorus and Potassium Fertilization Rates,” *Commun. Soil Sci. Plant Anal.*, vol. 49, no. 11, pp. 1314–1330, 2018, doi: 10.1080/00103624.2018.1457159.
- [8] S. Savci, “Investigation of Effect of Chemical Fertilizers on Environment,” *APCBEE Procedia*, vol. 1, no. January, pp. 287–292, 2012, doi: 10.1016/j.apcbec.2012.03.047.
- [9] A. Amkor, K. Maaider, and N. El Barbri, “Mint treatment day prediction using a multi-sensors system and machine learning algorithms,” *Sensors Actuators, A Phys.*, vol. 328, p. 112787, 2021, doi: 10.1016/j.sna.2021.112787.
- [10] A. Amkor, N. El Barbri, and K. Maaider, “A comparison between PLSR, SVMR and NARX network for the mint treatment day prediction based on multisensor system,” 2021 Int. Conf. Optim. Appl. ICOA 2021, 2021, doi: 10.1109/ICOA51614.2021.9442652.
- [11] A. Amkor and N. El Barbri, “A measurement prototype based on gas sensors for detection of pesticide residues in edible mint,” *J. Food Meas. Charact.*, 2020, doi: 10.1007/s11694-020-00617-8.
- [12] A. S. Petropoulos, A. Fernandes, N. Polyzos, V. Antoniadis, L. Barros, and C. F. R. I. Ferreira, “The Impact of Fertilization Regime on the Crop Performance and Chemical Composition of Potato,” *Agronomy*, vol. 10, pp. 1–18, 2020.
- [13] M. Yousaf, J. Li, J. Lu, T. Ren, R. Cong, S. Fahad and X. Li “Effects of fertilization on crop production and nutrient-supplying capacity under rice-oilseed rape rotation system,” *Sci. Rep.*, vol. 7, no. 1, pp. 1–9, 2017, doi: 10.1038/s41598-017-01412-0.
- [14] H. Zakir, M. Sultana, and K. Saha, “Influence of Commercially Available Organic vs Inorganic Fertilizers on Growth Yield and Quality of Carrot,” *J. Environ. Sci. Nat. Resour.*, vol. 5, no. 1, pp. 39–45, 2012, doi: 10.3329/jesnr.v5i1.11551.
- [15] T. Abedi, A. Alemzadeh, and S. A. Kazemeini, “Effect of organic and inorganic fertilizers on grain yield and protein banding pattern of wheat,” *Aust. J. Crop Sci.*, vol. 4, no. 6, pp. 384–389, 2010.
- [16] F. Yoldas, S. Ceylan, N. Mordogan, and B. C. Esetlili, “Effect of organic and inorganic fertilizers on yield and mineral content of onion (*Allium cepa* L.),” *African J. Biotechnol.*, vol. 10, no. 55, pp. 11488–11492, 2011, doi: 10.4314/ajb.v10i55.
- [17] I. Ugulu, K. Ahmad, Z. I. Khan, M. Munir, K. Wajid, and H. Bashir, “Effects of organic and chemical fertilizers on the growth, heavy metal/metalloid accumulation, and human

- health risk of wheat (*Triticum aestivum* L.),” *Environ. Sci. Pollut. Res.*, vol. 28, no. 10, pp. 12533–12545, 2021, doi: 10.1007/s11356-020-11271-4.
- [18] A. Sharma, “A Review on the Effect of Organic and Chemical Fertilizers on Plants,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. V, no. II, pp. 677–680, 2017, doi: 10.22214/ijraset.2017.2103.
- [19] Chandini, R. Kumar, R. Kumar, and O. Prakash, “The impact of chemical fertilizers on our environment and ecosystem,” *Res. Trends Environ. Sci.*, no. February, pp. 69–86, 2019.
- [20] F. Zhang, Y. Li, M. Yang, and W. Li, “Content of heavy metals in animal feeds and manures from farms of different scales in Northeast China,” *Int. J. Environ. Res. Public Health*, vol. 9, no. 8, pp. 2658–2668, 2012, doi: 10.3390/ijerph9082658.
- [21] C. Manyi-Loh, S. Mamphweli, E. Meyer, and A. Okoh, Antibiotic use in agriculture and its consequential resistance in environmental sources: Potential public health implications, vol. 23, no. 4. 2018.
- [22] K. Chester, S. Zahiruddin, A. Ahmad, W. Khan, S. Paliwal, and S. Ahmad, “Bioautography-based Identification of Antioxidant Metabolites of *Solanum nigrum* L. and Exploration Its Hepatoprotective Potential agChester, K. et al. (2017) ‘Bioautography-based Identification of Antioxidant Metabolites of *Solanum nigrum* L. and Explorati,” *Pharmacogn. Mag.*, vol. 13 (Suppl, no. 62, pp. 179–188, 2017, doi: 10.4103/pm.pm.
- [23] N. Wang, S. Chen, and Z. Zhou, “Age-dependent characterization of volatile organic compounds and age discrimination in Chinese rice wine using an untargeted GC/MS-based metabolomic approach,” *Food Chem.*, vol. 325, no. March, p. 126900, 2020, doi: 10.1016/j.foodchem.2020.126900.
- [24] M. T. Sánchez, A. Garrido-Varo, J. E. Guerrero, and D. Pérez-Marín, “NIRS technology for fast authentication of green asparagus grown under organic and conventional production systems,” *Postharvest Biol. Technol.*, vol. 85, pp. 116–123, 2013, doi: 10.1016/j.postharvbio.2013.05.008.
- [25] N. Jurado-Campos, M. García-Nicolás, M. Pastor-Belda, T. Bußmann, N. Arroyo-Manzanares, B. Jiménez, P. Viñas and L. Arce “Exploration of the potential of different analytical techniques to authenticate organic vs. conventional olives and olive oils from two varieties using untargeted fingerprinting approaches,” *Food Control*, vol. 124, no. August 2020, 2021, doi: 10.1016/j.foodcont.2020.107828.
- [26] C. Hernández-Sánchez, G. Luis, I. Moreno, A. Cameán, A. G. González, D. González-Weller, A. Castilla, A. Gutiérrez, C. Rubio and A. Hardisson, “Differentiation of mangoes (*Mangifera indica* L.) conventional and organically cultivated according to their mineral content by using support vector machines,” *Talanta*, vol. 97, pp. 325–330, 2012, doi: 10.1016/j.talanta.2012.04.038.
- [27] V. V. Merlini, F. D. L. Pena, D. T. Da Cunha, J. M. De Oliveira, M. A. Rostagno, and A. E. C. Antunes, “Microbiological Quality of Organic and Conventional Leafy Vegetables,” *J. Food Qual.*, vol. 2018, 2018, doi: 10.1155/2018/4908316.

- [28] C. Bigot, J. C. Meile, A. Kapitan, and D. Montet, "Discriminating organic and conventional foods by analysis of their microbial ecology: An application on fruits," *Food Control*, vol. 48, pp. 123–129, 2015, doi: 10.1016/j.foodcont.2014.03.035.
- [29] O. S. Papadopoulou, E. Z. Panagou, F. R. Mohareb, and G. J. E. Nychas, "Sensory and microbiological quality assessment of beef fillets using a portable electronic nose in tandem with support vector machine analysis," *Food Res. Int.*, vol. 50, no. 1, pp. 241–249, 2013, doi: 10.1016/j.foodres.2012.10.020.
- [30] N. El Barbri, E. Llobet, N. El Bari, X. Correig, and B. Bouchikhi, "Electronic nose based on metal oxide semiconductor sensors as an alternative technique for the spoilage classification of red meat," *Sensors*, vol. 8, no. 1, pp. 142–156, 2008, doi: 10.3390/s8010142.

CONCLUSION GENERALE

Les nez électroniques à base de capteurs de gaz sont désormais reconnus comme des alternatives fiables aux techniques de caractérisation traditionnelles des odeurs, permettant une plus large gamme d'applications dans les domaines de la sécurité alimentaire, de la surveillance environnementale et de la santé. L'objectif de nos travaux de recherche est la surveillance de la qualité des aliments largement consommés partout dans le monde.

À l'aide de capteurs de gaz à oxyde métallique semi-conducteur de commerce, nous avons conçu et réalisé un dispositif expérimental nez électronique polyvalent pour des utilisations de caractérisations des produits agricoles. Deux produits ont été étudiés : la menthe et les pommes de terre.

Les premiers tests d'analyse ont été menés sur la menthe comestible, cette substance est soumise à des traitements intensifs avec des pesticides de sorte insecticides qui mettent la vie humaine en danger nécessite une surveillance vigilante en raison de la forte demande. Pour la première application, nous avons démontré la capacité du dispositif conçu formé de 3 capteurs de gaz, deux de type TGS et un de type MQ, de discriminer trois types de menthe : menthe non traitée, menthe traitée avec l'insecticide Decis Fluxx et menthe traitée avec Malyphos 50. Après avoir prélevé des échantillons frais des trois classes préconnues, nous sommes parvenus à les bien trier selon leurs classes originales avec des taux de réussites très encourageants. En fait, l'utilisation de l'analyse en composantes principales (ACP) en combinaison avec la méthode des machines à vecteurs de support (SVM) a donné des taux de réussite entre 75% et 100% dans la discrimination des différents types de menthe.

Dans la deuxième application toujours dans le cadre de l'étude de la menthe, et grâce au même dispositif construit à base des dits trois capteurs, les composés volatils de la menthe traitée avec l'insecticide Malyphos 50 ont été étudiés pendant sept jours après le traitement. Il a été démontré qu'en utilisant trois caractéristiques de réponse des capteurs qu'il est possible d'identifier la menthe traitée et celle non traitée en utilisant la méthode des k-plus proches voisins (KNN) avec la validation croisée 5 fois qui a donné les résultats 97% de taux de réussite pour l'identification de la nature de la menthe. Ainsi, il a été validé qu'il est possible de prédire le jour de traitement de l'échantillon de menthe avec une grande précision en utilisant le réseau

autorégressif non linéaire avec entrée exogène (NARX) qui a réalisé un coefficient de corrélation de 0,99 et une erreur quadratique moyenne presque nulle $1,2986e-14$.

En ce qui concerne les deuxièmes tests d'analyse, ils ont été réservés aux pommes de terre, ce légume largement consommé dans le monde entier est aussi soumis à des traitements par des substances chimiques, ce qui rend sa consommation dangereuse pour la santé humaine. Cette fois notre dispositif a été basé sur cinq capteurs de gaz au lieu de trois que nous avons utilisé dans le cas de la menthe, trois de type TGS et deux de type MQ. Deux applications ont été réalisées.

Premièrement, nous avons étudié la discrimination entre deux types de pommes de terre selon leurs sources : champ traité avec les engrais NPK et champ non traitée. Pour atteindre cet objectif, des échantillons de pommes de terre ont été prélevés d'un champ non traité et d'autres échantillons d'un champ traité avec des engrais NPK sont soumis à notre processus d'analyse et les données résultantes ont été traitées par la suite. La méthode des machines à vecteurs de support (SVM) a été utilisée pour discriminer les données et elle a réalisé de meilleure performance avec une précision qui est arrivée à 92,5%.

Dans la deuxième application pour les pommes de terre, nous avons essayé d'élargir le champ de l'étude en attaquant trois types de pommes de terre selon leurs champs de production : traités traditionnellement avec le fumier de moutons et d'ânes, traités avec les engrais NPK et traités avec le fumier de poulet. L'analyse microbiologique a été utilisée comme méthode complémentaire dans cette étude pour démontrer la différence entre les échantillons de pommes de terre selon la nature des champs dans lesquels ils ont été récoltés, et une grande différence en termes de nombre de colonies bactériennes a été enregistrée entre les échantillons, en particulier après le quatorzième jour de l'essai, avec une petite différence enregistrée le septième jour qui a été exploité pour ouvrir la voie à l'analyse électronique du nez. Après avoir préparé les échantillons de pommes de terre de différents types selon la nature du traitement du champ, leurs informations (empreintes olfactives) ont été collectées à l'aide du même dispositif utilisé au paravent dans la précédente application qui est basé sur cinq capteurs de gaz de commerce. Ensuite, la méthode d'apprentissage automatique des machines à vecteurs de support (SVM) a été utilisée et elle a prouvé son efficacité à distinguer les sources de pommes de terre dans deux manières différentes : une méthode dite directe où toutes les données ont été utilisées à la fois qui a conduit à un taux de réussite de 92%. Une méthode composée comportant deux étapes : la première consiste à identifier les échantillons traités traditionnellement avec le fumier de

moutons et des ânes de ceux traités autrement, et la deuxième étape, l'identification de la nature traitement si c'est avec des engrais NPK ou du fumier de poulet, les résultats étaient de 96,7% pour la première étape et 92,5 % pour la seconde.

En guise de conclusion, le dispositif élaboré dans le cadre de ce travail de thèse s'est révélé efficace et a démontré à travers les nombreuses applications réalisées qu'il est capable d'être utilisé comme une bonne alternative par rapport aux différentes techniques physico-chimiques en termes de coût réduit, de portabilité et miniaturisation, de gain de temps et d'effort.

Pour les perspectives, le dispositif développé au cours de ce mémoire va être étendu à d'autres applications en agroalimentaire avec une amélioration. Les suggestions pour les travaux d'amélioration du dispositif peuvent être divisées en deux parties principales : la première partie sera consacrée à la technique de conception du nez électronique et comment il peut être développé davantage, et la deuxième partie sera consacrée à l'amélioration des techniques d'intelligence artificielle pour atteindre de très bons résultats.

RÉSUMÉ

L'identification des produits alimentaires, pollués et contaminés par des produits chimiques, ayant un impact négatif sur la santé humaine, est importante pour les consommateurs. Les travaux de recherches que nous rapportons dans cette thèse s'inscrivent dans le cadre du développement d'un dispositif « nez électronique portable », miniaturisé et à moindre coût, basé sur des capteurs de gaz orienté vers des applications visant à vérifier la qualité des produits alimentaires, notamment les produits traités avec des pesticides. Dernièrement, les nez électroniques ont été largement exploités dans différents domaines y compris l'agroalimentaire grâce à leur efficacité, leur précision, leur moindre coût et leur facilité d'utilisation, contrairement aux méthodes physico-chimiques. Les nez électroniques sont constitués de deux parties principales, une partie matérielle qui contient principalement un réseau de capteurs de gaz et une partie logicielle dédiée au traitement et classification de données.

Après une étude de marché sur les capteurs de gaz, notre choix s'est porté sur des capteurs de gaz à oxyde métallique du commerce qui seront utilisés pour la réalisation de notre dispositif. Dans un premier temps, une méthodologie de conception et d'impression de deux circuits est présentée, qui vise à optimiser le dimensionnement et l'agencement des capteurs d'un côté et à l'adaptation et la protection des composants d'un autre côté. Nous avons établi un protocole de mesure en se fixant des contraintes de sensibilité et de réponse. Un prototype avec une chambre en verre pour l'isolation, conçu d'après cette méthodologie, est testé, réalisé, mis en œuvre et caractérisé.

L'expérimentation de notre dispositif pour des échantillons de menthe traitée avec deux dangereux insecticides nous a permis de distinguer entre les échantillons avec des taux de réussites allant de 75% à 100% en utilisant la méthode des Séparateurs à Vaste Marges (SVM) dans un premier temps et K plus Proches Voisins (KNN) dans un autre temps et une prédiction de jour de traitement pour un insecticide avec un taux de corrélation qui est arrivé à 99% grâce à la méthode de Réseau Autorégressif non linéaire à entrées exogènes (NARX).

Ensuite, nous nous sommes intéressés à discriminer entre trois échantillons de pommes de terre selon leurs champs de culture : traité traditionnellement avec le fumier de moutons et d'ânes, traité avec des engrais NPK et traité avec le fumier de poulet. En appliquant la méthode SVM, nous avons atteint un taux de réussite qui dépasse 92%.

Au final, on peut dire que les nez électroniques ainsi développés se sont révélés être des outils de surveillance fiables de la contamination des produits alimentaires.

Mots clés : nez électronique, capteurs de gaz à oxyde métallique, discrimination entre la menthe traitée et non traitée, distinction des pommes de terre par le champ de culture, analyse de données, méthodes de reconnaissances de formes.