

La spectroscopie proche infra-rouge pour le phénotypage d'insectes

Myriam Siegart¹, Floriane Bouvier¹, Sandrine Maugin¹, Alain Lecomte¹, Claire Lavigne¹

Résumé. La détermination d'insectes est une discipline complexe mais pourtant indispensable si nous voulons évoluer vers une agriculture de précision. Actuellement, quelques experts sont en mesure d'identifier tous les insectes ravageurs de façon fiable, mais ils sont trop peu nombreux pour pouvoir répondre à toutes les demandes des agriculteurs. La mise au point d'une technique permettant de différencier sur le terrain des espèces de ravageurs serait donc très utile. Des techniques de biologie moléculaire existent déjà mais elles nécessitent une analyse en laboratoire qui peut être coûteuse. La spectroscopie proche infrarouge est une technologie utilisée depuis de nombreuses années dans d'autres disciplines. Elle présente la particularité d'être rapide (quelques secondes), non destructive, fiable, transportable sur le terrain et peu chère, une fois l'appareil acheté. Toute la difficulté de cette technique réside dans le fait d'avoir une base de données suffisamment robuste et à jour. Nous proposons ici des exemples d'utilisation de cette technique pour l'identification d'insectes. Un mode opératoire de l'acquisition de spectres mais aussi de l'analyse des données nécessaire à la construction d'un modèle prédictif est présenté. En conclusion, quelques exemples de résultats et d'applications sont présentés afin d'illustrer l'éventail du champ des possibles qu'offre cette technologie.

Mots clés : spectroscopie proche infra-rouge, phénotypage d'insectes, modèle prédictif, identification d'insectes


Introduction

La réduction de l'utilisation des pesticides et la mise en œuvre de méthodes de protection des cultures sont promues dans de nombreuses régions, pour diminuer les risques liés à l'utilisation des pesticides sur la santé des agriculteurs et sur la biodiversité. Un tel changement dans les stratégies de protection des cultures exige une meilleure connaissance des communautés d'organismes nuisibles car ces traitements demandent plus de précision pour être efficaces. Cette nécessité est d'autant plus importante dans les vergers qui reçoivent un grand nombre de traitements. En effet, les exigences de qualité des fruits liées au marché des produits frais sont telles que les seuils de tolérance en dégâts sont très bas.

Les pucerons et tordeuses sont les deux principaux groupes d'insectes nuisibles trouvés dans les vergers de pommiers. Parmi les tordeuses, deux espèces sont particulièrement fréquentes : la tordeuse orientale du pêcher (*Grapholita molesta* (Busck)), et le carpocapse des pommes (*Cydia pomonella* (Linné)). Leur abondance relative dépend des régions du monde. *G. molesta* est un ravageur important des vergers de pêcheurs en Europe. Ses larves peuvent causer des dommages importants à la fois sur les pousses et les fruits. Après la récolte, les adultes peuvent voler vers d'autres arbres fruitiers, tels que les cognassiers, les amandiers, les abricotiers, les pruniers et les cerisiers, où cette espèce est signalée comme un ravageur mineur (Balachowsky, 1966). Au cours des 20 dernières années, *G. molesta* a partiellement changé d'hôtes et est devenu capable de faire son cycle de vie complet sur des arbres fruitiers à pépins (Usmani and Shearer, 2001, Natale et al., 2003, Turata et al., 2005, Siegart et al., 2010). *C. pomonella*, une espèce proche de *G. molesta*, est un insecte ravageur majeur de fruits à pépins dans toutes les régions tempérées (Balachowsky, 1966).

Avec la récente co-occurrence des deux espèces dans les vergers de fruits à pépins, il est devenu important de mettre en place des programmes de surveillance, qui distinguent les proportions de ces deux espèces. L'identification des espèces le plus tôt possible dans la saison, ou même à la fin de la saison précédente, permet d'extrapoler les niveaux des populations qui seront présentes l'année suivante. Cela est utile d'abord pour cibler les traitements précoces (biologies différentes et donc ciblage/timing des traitements insecticides légèrement différents) et puis pour adapter le contrôle avec des méthodes ciblées (par exemple, les traitements avec un virus spécifique). Cependant, l'identification au stade larvaire est très complexe, d'une part à cause de la similitude des cycles de vie des deux espèces, d'autre part parce que les dégâts sur fruits sont également très similaires. De

1 Plantes et Systèmes de culture horticoles, INRA, 84000 Avignon, France
myriam.siegart@inra.fr



plus, les stades larvaires de ces deux espèces sont très difficiles à différencier (**Figure 1**). Certaines clefs de détermination morphologiques sont disponibles mais insuffisantes pour une identification précise, de plus cette observation nécessite l'utilisation d'une loupe binoculaire. L'identification des adultes est plus facile, car des piégeages peuvent être effectués au cours de la saison de vol à l'aide de pièges à phéromones spécifiques. Cependant, ces pièges ne sont pas très sensibles en raison des techniques de confusion sexuelle utilisées sur le terrain. Enfin, les méthodes moléculaires peuvent être utilisées pour distinguer ces deux espèces, en utilisant des marqueurs microsatellites comme (SST) (Franck et al., 2007, Chen et Dorn, 2010) ou les caractérisations moléculaires (« barcoding » basé sur le séquençage d'une portion du gène mitochondrial COI) (Huemer et al., 2012). Ces méthodes ont l'avantage d'être fortement spécifiques, mais nécessitent des équipements spéciaux coûteux.

Nous avons donc cherché à utiliser la technologie de spectroscopie proche infrarouge (SPIR) pour distinguer les larves de ces deux espèces de lépidoptères. La SPIR donne une image du profil chimique des échantillons en mesurant les vibrations d'élongation des liaisons chimiques entre atomes (Bertrand et Dufour, 2006). Cette technique a été longtemps utilisée dans l'industrie alimentaire (Kaffka et Czabaffy, 1981), et a eu plus récemment différentes applications en entomologie. Par exemple, elle a permis la différenciation de deux espèces de Noctuidae à des stades précoces : la tordeuse du tabac, *Heliothis virescens* (F.) et le ver de l'épi, *Helicoverpa zea* (Boddie) (Jia et al., 2007).

Matériel et méthodes

Échantillons d'insectes

Les insectes analysés étaient à leur cinquième et dernier stade larvaire. Il est possible de travailler sur des larves diapausantes piégées en verger commerciaux ou sur des larves actives issues d'élevage. Une des difficultés de ce genre de technique très précise est qu'il faut s'assurer que les stades physiologiques des insectes soient le plus homogènes possible. Pour cela, nous avons donc choisi dans les deux cas des larves qui avaient cessé de s'alimenter.

Les individus d'élevage ont été produits en masse sur un milieu artificiel fait maison à base de maïs (Guennelon, et al., 1981). Ces deux souches ont été élevées dans des chambres climatiques distinctes sous un cycle 16:8 h de lumière / obscurité, à 25°C, avec 70 % d'humidité relative. La procédure d'élevage est basée sur six générations par an pour le carpocapse et sept générations par an pour la tordeuse orientale. Les larves nouvellement écloses des deux espèces (âgées de 0 à 4 h) ont été placées individuellement dans des piluliers cubiques en plastique de petite taille (2 x 2 x 2 cm) contenant le milieu artificiel Stonefly *Heliothis* diet (WARD'S Natural Science®, Rochester, NY, USA) et incubées à 23°C pour les élever jusqu'au cinquième stade larvaire. Les populations sauvages sont recueillies à l'automne dans des vergers commerciaux. Nous avons délibérément choisi des individus issus de vergers différents pour avoir un niveau élevé de diversité génétique et biologique dans nos échantillons. Les larves ont été piégées avec du carton ondulé entouré autour des troncs (**Figure 1**). Les larves diapausantes y entrent pour passer l'hiver. Pour la réalisation de la base de données, les larves ont été identifiées morphologiquement par la personne responsable de l'élevage qui est experte pour ces insectes. La précision de cette identification a été vérifiée *a posteriori* à l'émergence des adultes. Elle s'est avérée sans erreur.



Figure 1. a) bandes pièges cartonnées pour la capture de larves diapausantes sur le terrain ; b) 5^e stade larvaire de *C. pomonella* ; c) 5^e stade larvaire de *G. molesta* (photos : M. Siegwart).

Procédures d'analyse

Spectroscopie proche infrarouge à transformée de Fourier

Les mesures sont faites à l'aide d'un spectromètre proche infrarouge à transformée de Fourier (MATRIX-F Transportable, BRUKER Optique®, Marne-la-Vallée, France). Lorsque l'appareil est au laboratoire, celui-ci ne doit pas être arrêté afin d'assurer le fonctionnement continu du dessiccateur qu'il contient, car les lentilles de cet appareil sont très sensibles aux variations d'humidité. La gamme spectrale de cet appareil s'étend de 780 nm à 2500 nm. La spécificité de cette spectroscopie est conférée par le guidage de la lumière infrarouge à travers un interféromètre. La lumière réfléchi sur l'échantillon est divisée en deux faisceaux identiques. L'un de ces faisceaux est réfléchi sur un miroir fixe et l'autre sur un miroir mobile avant d'atteindre le détecteur. Le détecteur reçoit l'interférogramme à toutes les longueurs d'onde. La fonction mathématique transformée de Fourier est utilisée pour convertir l'interférogramme en un spectre. En mode standard, ce spectromètre réalise 10 scans par échantillon et la moyenne de ces scans représente un spectre. La lumière émise par l'appareil est transmise au travers de fibres optiques d'une longueur de 3 mètres et terminées par une sonde en forme de pistolet (**Figure 2**). La lumière réfléchi de façon diffuse sur l'échantillon est captée par un détecteur puis transmise à travers d'autres fibres. Ces fibres optiques sont disposées aléatoirement à l'intérieur de cette sonde. Différents paramètres peuvent être réglés afin d'améliorer la qualité des spectres (le nombre de scans par échantillon, la résolution) ou pour réduire l'intensité lumineuse (filtres).

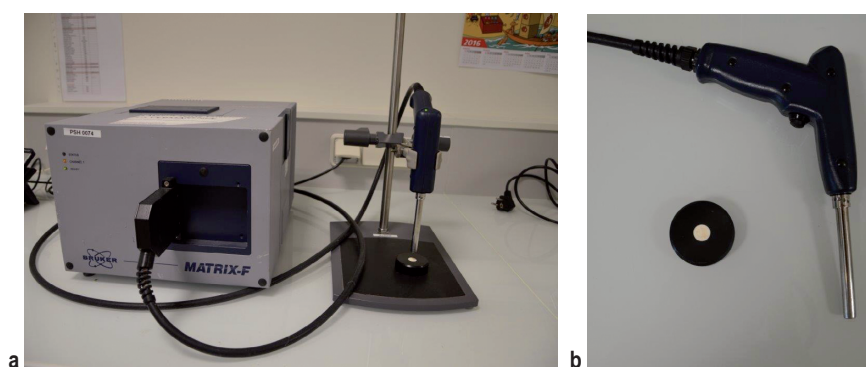



Figure 2. a) spectroscope proche infrarouge avec les accessoires nécessaires aux prises de spectres pour la détermination d'insectes ; b) sonde en forme de pistolet et spectralon® (photos : M. Siegwart).

Acquisition spectrale

Tout d'abord il faut activer la source de lumière infrarouge. Si le spectromètre a été arrêté avant cette utilisation, la mesure de l'interférogramme seule doit être réalisée avant de commencer les mesures. Ensuite l'appareil doit être étalonné en faisant une mesure de référence (blanc) à l'aide d'une plaque de réflexion diffuse : un spectralon® (1,5 cm de diamètre). Le pistolet est placé à 1 cm de cette surface et la lecture déclenchée. Cette surface doit





régulièrement être nettoyée à l'eau et à l'alcool avec du papier absorbant afin qu'elle conserve son rôle de « background ». Ce nettoyage peut également avoir lieu au cours d'un essai si le matériel scanné est salissant.

Les larves vivantes sont anesthésiées par le froid en les plaçant sur un pain de glace entouré d'une feuille de Sopalin® pendant 5 min pour garantir leur immobilité durant la mesure. Elles sont ensuite disposées sur leur face ventrale sur le spectralon®. Comme la longueur des larves dépasse le diamètre du faisceau lumineux (5 mm), seul le milieu du corps est scanné à l'aide de la sonde orientée verticalement à 10 mm au-dessus du spectralon®. Scanner un échantillon, incluant le positionnement de l'insecte et l'acquisition des spectres, prend moins d'une minute. Les spectres sont enregistrés grâce au logiciel d'acquisition de données OPUS 6.5 (Bruker) et sont convertis au format texte. Les données brutes sont composées d'une valeur d'intensité lumineuse pour chaque intervalle de 0,5 nm formant un spectre pour chaque individu. Ces données seront ci-après dénommées les spectres des individus.

Analyse des données

Aperçu

Des analyses statistiques sont effectuées dans le but de vérifier si la spectroscopie PIR permet une discrimination entre les espèces : *C. pomonella* et *G. molesta*. Dans un premier temps une analyse en composantes principales (ACP) permet de visualiser les données, puis un modèle prédictif est construit grâce à la méthode de PLS-DA (partial least square – discriminant analysis) afin de prédire l'appartenance d'un individu à un sous-ensemble. Les espèces des individus utilisés pour la construction du modèle doivent être connus. Nous utilisons le logiciel Matlab 7.5 (Mathworks Inc., Natick, MA) en utilisant le package SAISIR (Cordella et Bertrand, 2014) pour la construction de nos modèles. Les analyses de prétraitement spectrales sont faites avec R (R 2.13.2, packages 'car' et 'ade4' (R Development Core Team, 2011)).

Analyses de prétraitement et exploratoires

Tout d'abord, l'ensemble des spectres doit être inspecté à la recherche d'anomalies. Les spectres très différents qui sont dus à des mouvements de la larve durant la mesure sont éliminés (**Figure 3**). Les spectres sont ensuite normalisés pour les rendre comparables (Cordella et Bertrand, 2014). Une analyse de variance type ANOVA en utilisant le critère « espèce » comme facteur peut être réalisée sur l'ensemble de la gamme spectrale pour déterminer les zones spectrales les plus différentes entre les deux espèces testées (**Figure 4**). Par la suite, les analyses peuvent se faire uniquement sur ces zones. La méthode PLS-DA n'étant pas sensible au bruit de fond, cette étape n'est pas indispensable à la construction d'un modèle de prédiction. Cependant les ACP l'étant nettement plus, ce prétraitement est utile si l'on souhaite visualiser les données avec ce type d'analyse (**Figure 5**).

Etalonnage et évaluation des modèles PLS-DA

Les modèles PLS-DA sont construits (= calibration) à l'aide de 75 % des échantillons, les 25 % restant sont utilisés pour la validation croisée (= prédiction). Les échantillons sont aléatoirement placés dans une catégorie ou une autre. Ce tirage aléatoire suivi de la construction du modèle et sa validation sont répétés 10 fois, les résultats présentés correspondent à la moyenne de ces 10 répétitions. On attribue un score aux individus utilisés pour la construction du modèle en fonction de leur espèce (*C. pomonella* = 1; *G. molesta* = 2). Les dimensions du modèle (à savoir, le nombre de variables latentes retenues) varient de 1 à 10. La qualité du modèle résultant est évaluée une première fois par la valeur de son R^2 . La deuxième fois nous utilisons une validation croisée du modèle. C'est-à-dire que nous demandons au modèle de prédire l'espèce d'un individu connu et nous vérifions l'exactitude de la prédiction (proportions d'affectations correctes et fausses) (**Figure 6**). La règle suivante est utilisée pour la classification : si la valeur prédite pour un spectre varie de 0 à 1,5, la larve est classée comme *C. pomonella* ; si elle varie entre 1,5 et 3, la larve est classée comme *G. molesta*. Sinon, les larves ne sont pas assignées.

Exemples de résultats

Prétraitement

La **Figure 3** montre un exemple de prévisualisation des spectres pour élimination des spectres aberrants et traitement préliminaire. Comme on peut le voir, les différences entre spectres sont mineures et ne peuvent pas être analysées à l'œil nu.

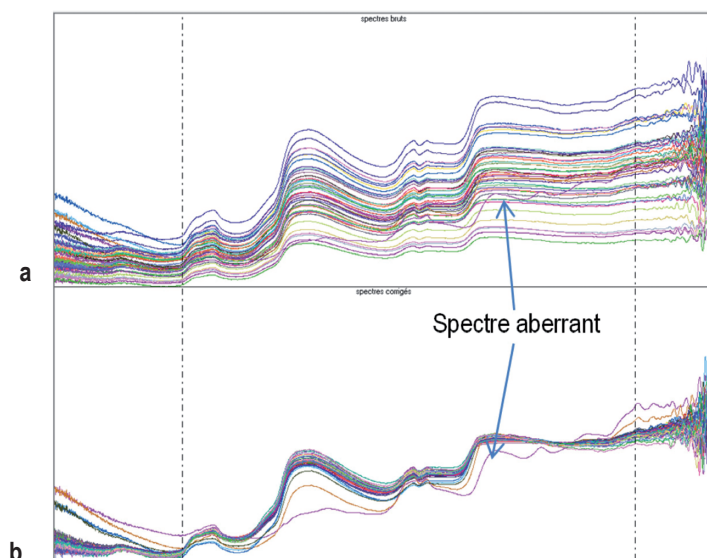


Figure 3. a) observation de l'ensemble des spectres (en absorbance) dans le proche infrarouge.

La gamme spectrale étudiée correspond au proche infrarouge (longueurs d'ondes comprises entre 800 et 2500 nm).

b) correction des spectres par la méthode SNV (standard normal variate). Elle permet de réduire les variations incontrôlées de la ligne de base. Ici le spectre TORD_Ly_CH_05_M_SF_D10_1 est aberrant que ce soit avant ou après traitement SNV. Il est donc retiré de la base pour la suite des analyses.

ANOVA et ACP

Une analyse de variance de type ANOVA peut être réalisée pour limiter la gamme spectrale analysée et ainsi limiter le bruit de fond.

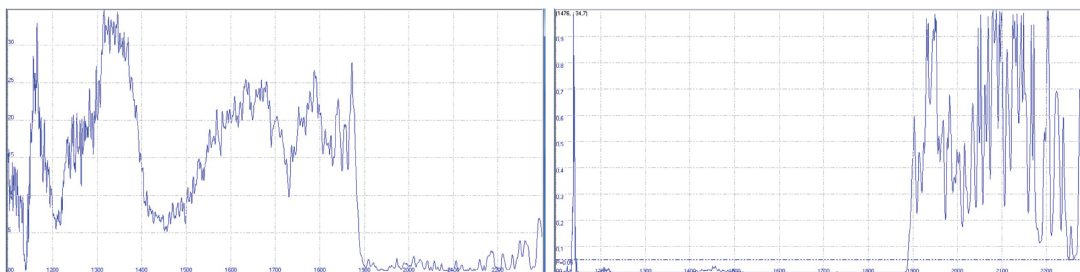


Figure 4. Le graphique de gauche donne la valeur de F (test F de Fisher) en ordonnée en fonction de la longueur d'onde en abscisse dans le cadre d'un test d'ANOVA à un facteur. Le facteur testé est l'alimentation des larves.

Le graphique de droite représente la probabilité (p) associée à l'hypothèse nulle (H0) réalisée dans le cadre du test d'ANOVA à un facteur. L'hypothèse nulle testée est la suivante : la composition générale des larves L5 femelles n'est pas affectée par leur alimentation larvaire (milieu artificiel à base de maïs vs milieu artificiel à base de soja).

Le seuil de 5 % est illustré par le trait horizontal en pointillé.



Ces graphiques montrent que dans la zone spectrale comprise entre 1 144,7 nm et 1 887,5 nm nous détectons une différence de composition générale des larves L5 femelles, significative en fonction du type d'alimentation.

La **Figure 5** montre une analyse en composante principale. Elle permet de visualiser les données et ainsi d'avoir une idée de la distance entre groupes d'échantillons, ici l'espèce des insectes. Dans cet exemple, les ellipses de confiance ne se chevauchent pas, ce qui laisse supposer une bonne discrimination des espèces d'insectes par la technique de spectroscopie proche infrarouge.

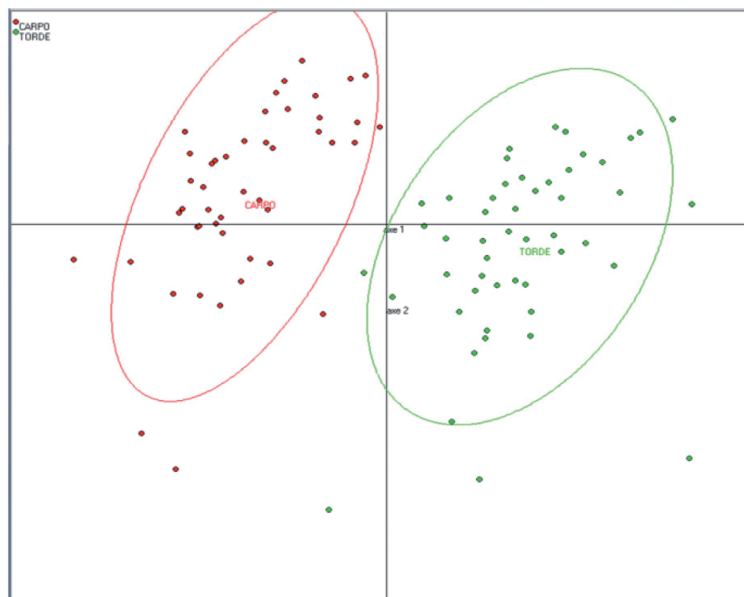
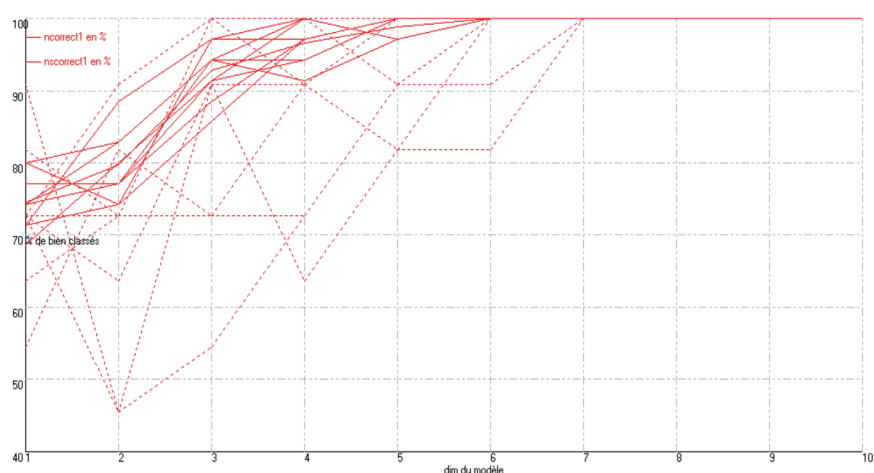


Figure 5. Analyse en composante principale de spectres proche infrarouge obtenus sur des larves de deux espèces de tortricidae. La gamme spectrale choisie ici est comprise entre 1130 et 2300 nm. 67,3 % de la variance est expliquée par l'axe 1 et 15,3 % par l'axe 2.

Evaluation des modèles PLS-DA

La **Figure 6** présente un exemple de validation d'un modèle PLS-DA visant à discriminer deux souches d'une même espèce d'insecte : *C. pomonella*.



	dim=1	dim=2	dim=3	dim=4	dim=5	dim=6	dim=7	dim=8	dim=9	dim=10
tst1	81,8%	72,7%	100,0%	100,0%	90,9%	100,0%	100,0%	100,0%	100,0%	100,0%
tst2	90,9%	45,5%	90,9%	63,6%	81,8%	100,0%	100,0%	100,0%	100,0%	100,0%
tst3	72,7%	45,5%	54,5%	72,7%	90,9%	100,0%	100,0%	100,0%	100,0%	100,0%
tst4	72,7%	90,9%	100,0%	90,9%	81,8%	100,0%	100,0%	100,0%	100,0%	100,0%
tst5	63,6%	72,7%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
tst6	54,5%	81,8%	72,7%	90,9%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
tst7	72,7%	90,9%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
tst8	72,7%	72,7%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%
tst9	81,8%	72,7%	72,7%	72,7%	90,9%	90,9%	100,0%	100,0%	100,0%	100,0%
tst10	72,7%	63,6%	90,9%	90,9%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Figure 6. Les courbes représentent le pourcentage d'individus bien classés (échelle des ordonnées) en fonction du nombre de dimensions de notre modèle PLS-DA (échelle des abscisses). Les courbes pleines traitent des individus utilisés pour la validation et les courbes pointillées ceux utilisés pour la calibration. Le tableau représente le pourcentage des individus (de la validation) bien classés pour chaque test (10 répétitions) en fonction du nombre de dimensions de notre modèle PLS-DA (partial least square).

Nous voyons ici qu'à partir de la dimension 7 aucun individu n'est mal classifié par notre modèle.

Conclusion

La SPIR est basée sur l'absorption du rayonnement infrarouge par l'échantillon analysé. Cette absorption est caractéristique des propriétés physico-chimiques des liaisons chimiques entre les atomes constitutifs de l'échantillon. Une analyse statistique multivariée des spectres d'absorption ou de transmission (à savoir, la description de l'intensité lumineuse en fonction de la longueur d'onde) permet la caractérisation ou la classification des échantillons. Cette technique, maintenant bien établie présente plusieurs avantages. Elle est rapide, transportable sur le terrain, facile à réaliser, automatisée, non destructive et peu coûteuse (après l'achat de l'équipement). Elle permet la caractérisation d'un grand nombre d'échantillons (phénotypage à haut débit) (Jia et al., 2007, Antonialli et al., 2008, Newey et al., 2008). Toutefois, elle est moins spécifique que les technologies moléculaires ou chimiques telles que chromatographie en phase gazeuse et spectrométrie de masse (Newey et al., 2008) et sa fiabilité dépend de la base de données initiale utilisée par le modèle de prévision. Plus celle-ci sera grande, plus le modèle sera robuste.

Cette technologie peut permettre d'identifier des larves de *C. pomonella* et *G. molesta* rapidement et avec précision, et même d'aller au-delà d'une discrimination inter-espèce. Elle permet également de discriminer des larves d'une même espèce en fonction de leur alimentation, leur âge ou bien leur souche.



Références bibliographiques

Antoniali WF, Suarez YR, Izida T, Andrade LHC, Lima SM (2008) Intra- and interspecific variation of cuticular hydrocarbon composition in two *Ectatomma* species (Hymenoptera: Formicidae) based on Fourier transform infrared photoacoustic spectroscopy. *Genet Mol Res* **7** : 559-566.

Balachowsky AS (1966) Entomologie appliquée à l'agriculture. Vol. 2. Masson et Cie, Paris.

Bertrand D, Dufour E (2006) La spectroscopie infrarouge et ses applications analytiques. Lavoisier, Paris.

Chen MH, Dorn S (2010) Cross-amplification of microsatellites from the codling moth *Cydia pomonella* to three other species of the tribe Grapholitini (Lepidoptera: Tortricidae). *Mol Ecol Res* **10** : 1034-1037.

Cordella CBY, Bertrand D (2014) SAISIR : A new general chemometric toolbox. *Trac-Trends Anal Chem* **54** : 75-82.

Franck P, Reyes M, Olivares J Sauphanor B (2007) Genetic architecture in codling moth populations: comparison between microsatellite and insecticide resistance markers. *Mol Ecol* **16** : 3554-3564.

Guennelon G, Audemard H, Fremont JC, Ammari ME (1981) Permanent rearing of codling moth on an artificial medium. *Agronomie* **1** : 59-64.

Huemer P, Zlatkov B, Baixeras J (2012) *Dichrorampha dinarica*, new species, a century of confusion in European lepidopterology (Lepidoptera: Tortricidae) resolved by combining morphology and DNA barcoding. *Zootaxa* **3389** : 41-50.

Jia FY, Maghirang E, Dowell F, Abel C, Ramaswamy S (2007) Differentiating tobacco budworm and corn earworm using near-infrared spectroscopy. *J Econ Entomol* **100** : 759-764.

Kaffka KJ, Czabaffy A (1981) The correlation between quality parameters and optical transmittance of some stone fruits determined with a near-infrared composition analyser. *Acta Alimentaria* **10** : 11.

Natale D, Mattiacci L, Hern A, Pasqualini E, Dorn S (2003) Response of female *Cydia molesta* (Lepidoptera : Tortricidae) to plant derived volatiles. *Bull Entomol Res* **93** : 335-342.

Siegwart M, Coupard H, Mottet C, Sauphanor B (2010) Recrudescence de la tordeuse orientale. *Phytoma* **633** : 28-32.

Newey PS, Robson SKA, Crozier RH (2008) Near-infrared spectroscopy as a tool in behavioural ecology: a case study of the weaver ant, *Oecophylla smaragdina*. *Anim Behav* **76** : 1727-1733.

Turata R, Mori N, Girolami V (2005) Adaptation of *Cydia molesta* (Busck) to pome fruit. *Inf. Fitopatol.* **55** : 45-49.

Usmani KA, Shearer PW (2001) Susceptibility of male Oriental fruit moth (Lepidoptera : Tortricidae) populations from New Jersey apple orchards to azinphosmethyl. *J Econ Entomol* **94** : 233-239.