

CHIMIOMÉTRIE 2016

17-20 Janvier 2016, Namur, Belgique

Comparaison des performances des modèles d'analyse discriminante PLS et SVM appliqués sur des spectres proche infrarouge de racines et de pailles de froment d'hiver.

Comparison of PLS and SVM discriminant analysis models performances when applied on near infrared spectra of roots and crop residues of winter wheat.

Damien Eylenbosch¹, Juan Antonio Fernández Pierna², Vincent Baeten², Bernard Bodson¹

¹ Département AGROBIOCHEM, Unité de Phytotechnie tempérée
Gembloux Agro-Bio Tech, Université de Liège
Passage des déportés 2, 5030 Gembloux, Belgique
e-mail : d.eylenbosch@ulg.ac.be

² Département Valorisation des productions, Unité Qualité des produits
Centre wallon de Recherches agronomiques
Chaussée de Namur, 15, 5030 Gembloux, Belgique
e-mail : j.fernandez@cra.wallonie.be

L'imagerie hyperspectrale proche infrarouge est de plus en plus utilisée pour analyser des produits agricoles. Cette méthode qui combine la spectroscopie proche infrarouge et l'imagerie permet d'enregistrer, pour chaque pixel de l'image, un spectre comprenant la réflectance à différentes longueur d'ondes. Cette analyse permet d'appréhender l'hétérogénéité et la variabilité spectrale des échantillons [1]. Des algorithmes adéquats sont nécessaires pour analyser les larges volumes de données générés.

Les modèles d'analyses discriminantes PLS (*Partial Least Squares*) et SVM

(*Support Vector Machines*) ont été comparés dans une étude qui avait pour objectif de discriminer des spectres de racines et de pailles de froment en vue de les quantifier. L'utilisation de l'imagerie hyperspectrale est proposée comme une alternative rapide et fiable au tri manuel qui est réalisé dans le cadre d'études sur les systèmes racinaires [2]. Suite au lavage des échantillons de sol, les racines et résidus de culture qui sont récupérés sont déposés sur une bande transporteuse passant sous une caméra hyperspectrale de type « line-scan » [3]. Les spectres acquis

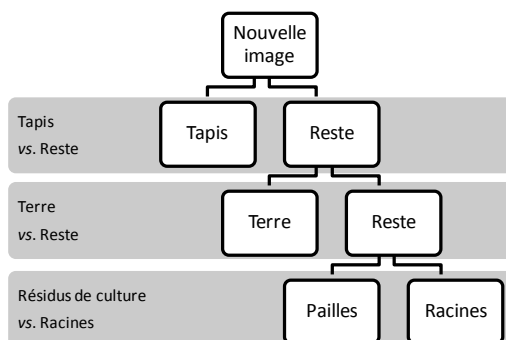


Figure 1: Arbre de décision dichotomique

sont ensuite analysés grâce aux algorithmes d'analyses discriminantes afin de prédire la nature de chaque pixel de l'image. La quantification des différents éléments est ensuite basée sur le nombre de pixels attribués à chaque classe. Dans le cadre de cette étude, les modèles permettaient aussi de prendre en charge les spectres proches infrarouges du tapis de mesure défilant sous la caméra et de la terre qui peut encore se trouver en faible quantité sur les échantillons.

Les modèles ont été construits à l'aide de la PLS Toolbox 7.8 (Eigenvector Research, Inc., Wenatchee, WA, USA) fonctionnant sur Matlab (The MathWorks, Inc., Natick, MA, USA). Pour la calibration et la validation des modèles, la méthode Duplex a été utilisée afin de sélectionner des spectres représentatifs de la variabilité spectrale des échantillons [4]. Un prétraitement des spectres se limitant à un centrage des moyennes suivi d'une division par l'écart-type (autoscale) a donné les meilleurs résultats. Pour la prédiction sur images, trois modèles ont été combinés au sein d'un arbre de décision dichotomique afin d'extraire successivement les spectres du tapis de mesure, de la terre, des pailles et des racines (Fig. 1).

Les résultats obtenus ont mis en évidence les meilleures performances des modèles SVM, aussi bien lors de la calibration que lors de la validation, et quelle que soit la classe prédite. Ces modèles ont permis de prédire correctement la nature d'au moins 85 % des spectres. Dans le cas des spectres infrarouge du tapis et de la terre, plus de 97 % des spectres étaient correctement classés. La classification était légèrement moins bonne pour les racines et les pailles. Cette observation est particulièrement vraie pour les spectres des pixels en bordure des pailles, ce qui semble être dû aux zones d'ombres qui perturbent l'acquisition et la discrimination des spectres (Fig. 2). Lors de la calibration du modèle permettant de discriminer les pailles et les racines, l'utilisation d'un modèle SVM a permis une amélioration des prédictions de 9 % par rapport à un modèle PLS.

Dans cette étude, les modèles SVM ont donc démontré leurs plus grandes performances et ont permis de mettre au point une méthode de discrimination rapide et fiable des racines et des résidus de culture de froment qui est utilisable dans l'étude de systèmes racinaires par prélèvement de carottes de sol aux champs.

Références

- [1] J. A. Fernández Pierna, V. Baeten, J. Dubois, J. Burger, E. N. Lewis & P. Dardenne, «NIR Imaging – Theory and applications». In: Comprehensive Chemometrics, Oxford, Elsevier (S. Brown, R. Tauler and B. Walczak eds.), Vol 4, pp. 173-196, 2009.
- [2] D. Eylembosch, J. A. Fernández Pierna, V. Baeten, B. Bodson, «Detection of wheat root and straw in soil by use of NIR hyperspectral imaging spectroscopy and Partial Least Square discriminant analysis». In: proceedings of the ESA XIIIth Congress, Debrecen, Hungary, pp. 237, 2014.
- [3] P. Vermeulen, J. A. Fernández Pierna, H. P. van Egmond, P. Dardenne, V. Baeten, «Online detection and quantification of ergot bodies in cereals using near infrared hyperspectral imaging». Food Addit Contam A, Vol 29, pp. 232, 2012.
- [4] R. D. Snee. «Validation of regression models: methods and examples». Technometrics, Vol 19, pp. 415, 1977.



Figure 2: Résultat de la prédiction sur un échantillon de paille