



Research Article

Volume 25 Issue 2 - October 2020  
DOI: 10.19080/ARTOAJ.2020.25.556296

Agri Res & Tech: Open Access J

Copyright © All rights are reserved by Cédric Camps

# FT-NIR assisted Machine and Deep learning for determination of Acteoside, Aucubin and Catalpol contents of *Plantago lanceolata*



Cedric Camps<sup>1\*</sup>, Xavier Simonnet<sup>1</sup>, Valérie Bardot<sup>2</sup> and Michel Dubourdeaux<sup>2</sup>

<sup>1</sup>Agroscope, Institute for Plant Production Sciences IPS, Switzerland

<sup>2</sup>Pileje Industrie, Naturopôle Nutrition Santé, Les Tiolans, F-03800 Saint-Bonnet-de-Rochefort, France

**Submission:** September 30, 2020; **Published:** October 05, 2020

\***Corresponding author:** Cédric Camps, Agroscope, Institute for Plant Production Sciences IPS, Route des Eterpys 18, 1964 Contthey, Switzerland

## Abstract

Aucubin, acteoside and catalpol are important compounds present in this plant but their determination is time consuming, costly and necessitate the use of chemicals. In the present study we developed a method based on FT-NIR spectroscopy to determine aucubin, acteoside and catalpol from dry powders of *Plantago lanceolata*. FT-NIR spectra were processed in PLS and deep learning methods and the accuracies of models were calculated and compared. Aucubin, acteoside and catalpol contents were predicted with a root mean square error of 0.56, 0.25 and 0.18%, respectively by using PLS method. Deep learning did not allow to improve the accuracies but allowed obtaining similar accuracies. FT-NIR based method showed interesting results and is promising to develop a fully usable method for selection programs of plantain with high level contents of bioactive compounds.

**Keywords:** Plantago Lanceolate, FT-NIR Spectroscopy, PLS, Deep Learning, aucubin, acteoside, Catalpol

## Introduction

*Plantago lanceolata* L. is a common grassland and roadside plant, traditionally used for animal and human food [1-3] and for its therapeutic properties [4]. *Plantago lanceolata* leaves are used in many European countries for the symptomatic treatment of colds and of mouth and throat inflammation [4]. Pharmacological investigations performed with the total leaf extract and with isolated agents from the total extract have reported anti-inflammatory, antioxidant, immunostimulant, epithelising, antitoxic, procoagulant, and spasmolytic effects as well as antibacterial, antiviral, and anthelmintic properties [4]. The iridoid glycosides, aucubin and catalpol and the phenylethanoid glycoside, acteoside are amongst the main bioactive components of *Plantago lanceolata* leaves [4]. The content in bioactive compounds in *Plantago lanceolata* is genetically determined but it can also be affected by various environmental or agronomic factors such as plant density or relative humidity [5,6].

Bioactive components are currently quantified by expensive analytical methods that necessitate the use of chemicals, time, high competences and know-how. High performance liquid

chromatography [HPLC] and high performance thin-layer chromatography [HPTLC] are often used to determine bioactive in samples in the framework of variety selection programs or to control new varieties [7,8]. Variety selection programs need a rapid and low-cost method to determine bioactive components in series of thousands of samples. FT-NIR [Fourier-transform Near-Infrared-Spectroscopy] supported by chemometric analysis could be an asset to reduce the time and cost of such analyses. Such a method has been successfully developed to quantify bioactive compounds in other plant species [9,10]. Thus, the present paper presents the possibility of measuring bioactive compounds in different varieties of *Plantago lanceolata* by FT-NIR and chemometrics.

## Material and Methods

### Plant material

Five varieties of *Plantago lanceolata* have been used in the present experiment (Table 1). Two were from Germany, one from New-Zealand, one from Canada and one from Switzerland.

**Table 1:** Varieties of *Plantago lanceolata* and their origins.

| Code | Variety  | Country | Supplier            |
|------|----------|---------|---------------------|
| PL1  | Noflor   | CH      | Mediseeds           |
| PL2  | UnKnown  | CA      | Richters            |
| PL3  | Libor    | DE      | Pharmasaat          |
| PL4  | Arterner | DE      | Pharmasaat          |
| PL5  | KRC7877  | NZ      | PGG Wrightson seeds |

### HPTLC analyses

**Extraction:** 500mg of dried and crushed leaves were taken and added with 45mL of a water-ethanol solution (50-50). The mixture was homogenized with a vortex for 10s, placed in an ultrasonic bath for 10 min at 60°C and homogenized again using a vortex for 10s. The homogenate was filtered and the filtrate collected in a 50mL flask using 5 ml of the water-ethanol solution (50-50). The flask was completed to 50 ml with the water-ethanol solution (50-50). 1mL was taken from this solution for HPTLC analysis.

**Phenylpropanoids:** Standards of acteoside, aucubin and catalpol are prepared at a concentration of 1 mg / ml. A range of 5 volumes of each standard solution [12µL, 9µL, 7µL, 4µL, 2µL] is deposited for HPTLC analysis. For the analysis of acteoside the migration solution used is composed of ethyl acetate [25mL], ultrapure water [6.75mL], formic acid [2.75mL] and acetic acid [2.75mL]. 4 mL were applied on the HPTLC plate. The silica plate was scanned at 330 nm for determination of acteoside concentration.

**Iridoids:** For quantitation of aucubin and catalpol, the migration solution was composed of chloroform (35mL), methanol (20mL) and 5mL of a solution of ammonia (1mL of trifluoroacetic acid in 50mL of ammonia at 25%). A derivatization reagent was used (10% sulfuric acid in methanol) in which the silica plate was dipped at a speed of 5m/sec for 1s. The plate was then heated at

110 °C for 3 min and then scanned at 450 nm for the estimation of concentrations of aucubin and catalpol.

### FT-NIR spectroscopy

#### Chemometric

**Partial least square regression:** Spectra were gathered in a matrix  $X_{n,p}$  where  $n$  is the number of spectra and  $p$  the number of wavenumber steps. The reference-values (Aucubin, Acteoside and Catalpol) were gathered in column vectors  $y_{n,1}$ . Samples were separated in a calibration set and a validation set. The accuracy and goodness of models has been evaluated according to several indicators: the coefficient of determination ( $R^2$ ), root mean square errors (RMSE), the ratio of prediction to deviation (RPD) [11] and the range error ratio (RER). The values of the beta-coefficient of the first latent variable were plotted to evaluate the weight of wavelength absorbance of each model. All data analyses were performed with OPUS software and Matlab R2019b (The MathWorks, Inc., Natick, MA, USA).

#### Deep learning

Deep learning approach has been performed using the “ntraintool” of Matlab R2019b® (The MathWorks, Inc., Natick, MA, USA). To predict the values of chemicals, the trainlm algorithm which is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization has been computed. Then, data were divided into a calibration data set and a validation data set for each chemical compound to be predict (Table 2). The results were summarized on the basis of correlation values of calibration and validation steps ( $R_{cal}$  and  $R_{val}$ ), the root mean square error of calibration and validation (RMSEC and RMSECV), the ratio performance to deviation (RPD), the ratio range to deviation (RER), the overall presentation of spectral data, the plot of actual versus predicted values of the chemical contents and the plot of residuals regardless of the standard deviation of the chemical contents values.

**Table 2:** PLS-values of aucubin, acteoside and catalpol.

| PLS Parameters           |         | Aucubin                     | Acteoside                     | Catalpol                      |
|--------------------------|---------|-----------------------------|-------------------------------|-------------------------------|
| n                        | CAL/VAL | 192/48                      | 193/55                        | 155/25                        |
| Wavenumber ranges (nm)   | CAL/VAL | 1332-1836                   | 1332-1836<br>2172-2357        | 1332-2175                     |
| LV                       | CAL/VAL | 11                          | 5                             | 15                            |
| Corr. Coef. (R)          | CAL/VAL | 0.88/0.88                   | 0.73/0.77                     | 0.80/0.65                     |
| RMSE                     | CAL/VAL | 0.54/0.56                   | 0.26/0.25                     | 0.19/0.18                     |
| RPD                      | CAL/VAL | 2.08/2.15                   | 1.47/1.61                     | 1.41/1.51                     |
| RER                      | CAL/VAL | 9.11/9.16                   | 6.85/7.32                     | 8.05/6.50                     |
| Offset                   | CAL/VAL | 0.55/0.75                   | 0.63/0.67                     | 0.30/0.35                     |
| slope                    | CAL/VAL | 0.78/0.69                   | 0.53/0.53                     | 0.55/0.47                     |
| Pre-treatment of spectra |         | MSC                         | D1(17) + SS                   | D1(17)                        |
| Content (% FW)           | CAL/VAL | [0.6 - 5.52] / [0.4 - 5.53] | [0.66 - 2.44] / [0.88 - 2.71] | [0.08 - 1.61] / [0.28 - 1.45] |

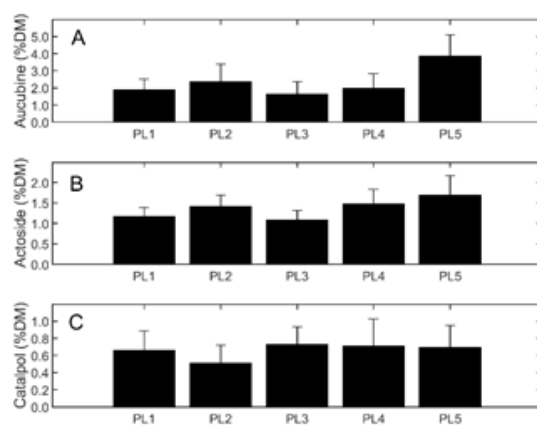
CAL: calibration, VAL: validation, n: number of spectra samples, LV: the number of latent variables, R: correlation coefficient, RMSE: root mean square error, RPD: ratio performance to deviation, RER: range error ratio. MSC: multi-scatter corrections, SS: straight line subtraction, D1(n): first derivative with a stepwise of 17.

## Results and discussion

### Chemical composition

Aucubin content ranged from less than 1% for the first variety to more than 5% for the fifth variety (Figure 1). The content of

the three other varieties being intermediate. Acteoside content ranged from 0.5% to 2.6% with the highest values for the fifth variety. Finally, catalpol content was generally very low, under 1% and the levels quite similar for all varieties.

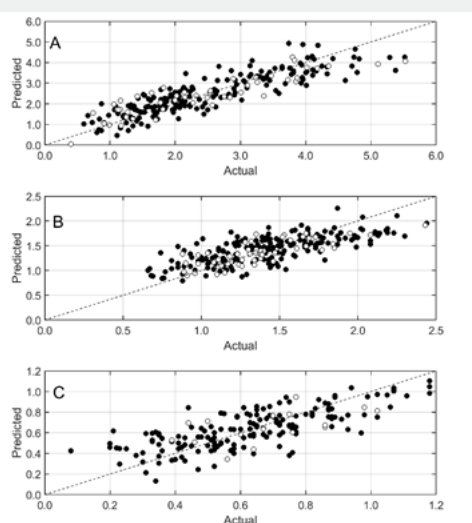


**Figure 1:** Aucubin (A), Acteoside (B) and Catalpol (C) content of the 5 *Plantago lanceolata* varieties (PL1-5).

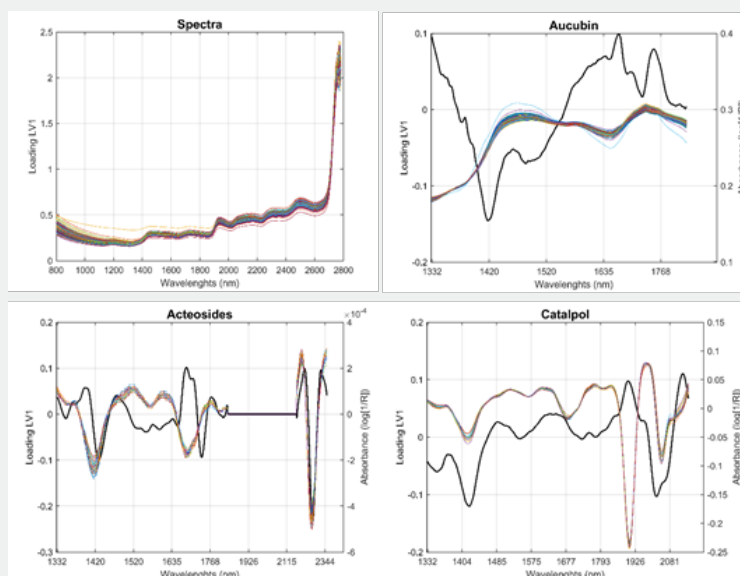
### PLS prediction of chemical composition

Partial least square regressions (PLS) were carried out to model the prediction of aucubin, acteoside and catalpol by FT-NIR spectroscopy. Table 2 gathers the statistics of the models. Aucubin and acteoside contents were the most accurately predicted with R-values of 0.88 and 0.77, respectively. The accuracies of these two models were around 0.56 %DW and 0.25%DW, respectively. Such results are visually confirmed by the plot of actual vs predicted values (Figure 2A and 2B). The models were built on the wavelength region 1332-1836nm for both models plus 2175-

2357nm for acteoside model. Loading-values of both models were plotted on figure 3. Prediction models of aucubin and acteoside relied on 2 common absorption bands in the vicinity of 1420nm and 1660-1680nm. Prediction of catalpol content was less accurate with root mean square error around 0.2%DW and correlation coefficient in validation step of 0.65 (Table 2). As expected, the very low range of catalpol values, sometimes less than 0.1%, did not allow a correct prediction. The model relied on one common wavelength absorbance around 1420nm but other parts of the spectra were less relevant in the prediction compared to the two other models (aucubin and acteoside) (Figure 3).



**Figure 2:** Scatter plot of actual vs PLS-predicted values of chemical contents of plantain leaves. Calibration (•), validation (○). A: aucubin, B: Acteosides, C: Catalpol. Units are expressed in % of Dry matter.



**Figure 3:** FT-NIR data collection, pre-processed FT-NIR spectra (colored lines) and beta-coefficients of the first latent variable (Black line) of the PLS regression models for acteoside, aucubin and catalpol predictions.

## Deep learning

An approach based on deep-learning analyses was attempted with the aim of improving the results obtained with the partial least square method. In this approach, the data set divided into a calibration and a validation set, as previously and the entire spectra were processed in the analyses. Table 3 shows the values of the models for each chemical content. The accuracy of acteoside and aucubin models remained correct with root mean square error of 0.24%DW and 0.47%DW. RPD-values show correct levels for selection programs but values too low for analytical prediction. RER-values were promising and showed that overall range of

chemical values was suitable for modelling. Figure 4 presents the plots issued from the deep-learning analysis for the prediction of aucubin content. However, the analysis of the residues showed that several samples were outside the limit values set at  $\pm 2SD$  (Figure 4C). On this basis, a few spectra downgraded as outliers. Finally, Figure 4B expresses the predicted values as a function of the measured values. This plot confirmed the promising results given by the relatively low values of quadratic error (RMSECV = 0.47). Similar deep learning approach was conducted for determination of acteosides and catalpol contents and determination of outliers (plots not shown).

**Table 3:** Deep learning results of acteoside, aucubin and catalpol contents predictions.

| Compounds  | n   | Rcal | RMSEC | Rval | RMSECV | RPDcv | RERcv |
|------------|-----|------|-------|------|--------|-------|-------|
| Acteosides | 248 | 0.73 | 0.13  | 0.8  | 0.24   | 1.67  | 7.63  |
| Aucubin    | 240 | 0.88 | 0.07  | 0.89 | 0.47   | 2.6   | 10.91 |
| Catalpol   | 181 | 0.74 | 0.17  | 0.72 | 0.24   | 1.08  | 4.88  |

## Conclusion

The present study aimed at evaluating the possibility to develop a rapid and non-invasive method to determine the aucubin, acteoside and catalpol contents of *Plantago lanceolata*. The interest of this approach is to be able to determine the chemical compounds of several hundreds or thousands of samples in a short period of time and at a low cost. On the basis of FT-NIR spectra obtained, the prediction levels were correct either with one (PLS) or the other method (Deep learning) of data analysis. Consequently, deep learning approach did not allow to significantly improve the prediction accuracies. Classic PLS-regression would be the method to be recommended to determine

aucubin, acteoside and catalpol contents. In further trials, the models will be improved by adding more variability by including more varieties of *Plantago lanceolata*.

## Acknowledgment

We thank the company PGG Wrightson seeds (NZ) for the provision of seeds of *Plantago lanceolata* cv. KRC7877.

## References

1. W Rumball, RG Keogh, GE Lane, JE Miller, RB Claydon (1997) Grasslands Lancelot' plantain (*Plantago lanceolata* L.). New Zealand Journal of Agricultural Research 40 : 373-377.

2. PM Guarrera, V Savo (2016) Wild food plants used in traditional vegetable mixtures in Italy. *Journal of Ethnopharmacology* 185 : 202-234.
3. TJ Fraser, SM Scott, JS Rowarth (1996) Pasture species effects on carcass and meat quality. *Proceedings of the New Zealand Grassland Association* 58 : 63-66.
4. S Navarrete, PD Kemp, SJ Pain, PJ Back (2016) Bioactive compounds, aucubin and acteoside, in plantain (*Plantago lanceolata* L.) and their effect on in vitro rumen fermentation. *Animal Feed Science and Technology* 222 : 158-167.
5. [https://www.ema.europa.eu/en/documents/herbalreport/final-assessment-report-plantago-lanceolata-l-folium\\_en.pdf](https://www.ema.europa.eu/en/documents/herbalreport/final-assessment-report-plantago-lanceolata-l-folium_en.pdf)
6. E Osinska, W Roslon, A Geszprych, A Morawska, L Gontar (2014) Effect of plant density on growth, development and quality of Ribwort Plantain (*Plantago Lanceolata* L.). *Journal of International Scientific Publications: Agriculture and Food 2* : 267-277.
7. S Gonda (2012) Effect of High Relative Humidity on Dried *Plantago lanceolata* L. Leaves during Long-term Storage: Effects on Chemical Composition, Colour and Microbiological Quality. *Phytochemical Analysis* 23: 88-93.
8. M Al-Mamun, D ABE, H KOFUJITA, Y TAMURA, H SANO (2008) Comparison of the bioactive components of the ecotypes and cultivars of plantain (*Plantago lanceolata* L.) herbs. *Animal Science Journal* 79 : 83-88.
9. C Camps (2014) Prediction of essential oil content of oregano by hand-held and Fourier transform NIR spectroscopy. *Journal of the Science of Food and Agriculture* 94: 1397-1402.
10. C Camps, M Toussirot, M Quennoz, X Simonnetb (2011) Determination of artemisinin and moisture content of *Artemisia annua* L. dry powder using a hand-held near infrared spectroscopy device. *Journal of Near Infrared Spectroscopy* 19 : 191-198.
11. P Williams, D Sobering (1993) Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *Journal of Near Infrared Spectroscopy* 1: 25-32.



This work is licensed under Creative Commons Attribution 4.0 License  
DOI: [10.19080/ARTOAJ.2020.25.556296](https://doi.org/10.19080/ARTOAJ.2020.25.556296)

### Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats  
( Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission  
<https://juniperpublishers.com/online-submission.php>