Comparison between single and mixed-species NIRS databases’ accuracy of dairy fiber feed value detection

I Agustiyani¹, Despal²*, L A Sari³, R Chandra³, R Zahera², I G Permana²

¹Study Program Nutrition and Feed Science, Department of Nutrition and Feed Technology, Faculty of Animal Science, IPB University, Jl. Agatis Kampus IPB Darmaga 16168, Bogor-Indonesia
²Department of Nutrition and Feed Technology, Faculty of Animal Science, IPB University, Jl. Agatis Kampus IPB Darmaga 16168, Bogor-Indonesia
³Study Program Nutrition and Feed Technology, Department of Nutrition and Feed Technology, Faculty of Animal Science, IPB University, Jl. Agatis Kampus IPB Darmaga 16168, Bogor-Indonesia

*E-mail: despaltk@gmail.com

Abstract. Near Infrared Reflectance Spectroscopy (NIRS) accuracy is affected by its database. However, our previously developed database for dairy cattle dietary fiber feed (DFF) showed low accuracy for complex organic substance detection due to mixed-species used in the database. This paper aimed to compare single and mixed-species NIRS database accuracy in predicting DFF nutrient contents. In the mixed database, five feeds from three areas of dairy cattle farming were sampled. In the single database, thirty Napier grass from 30 areas were collected. Samples were analyzed chemometrically for their nutrient contents. Spectra of each sample were collected three times (two spectra for calibration and a spectrum for validation) using FT-NIRS Spectrometer Solid Cell. Calibration and validation models were carried out using the Partial Least Squares (PLS) regression. For external validation, seven independent Napier grass samples were tested. The result showed that the single species NIRS database developed using Napier grass was less accurate than mixed-species DFF due to huge nutrient content variations between varieties of Napier grass. It is concluded that database accuracy developed from mixed dietary fiber feed were more accurate in comparison to single species and suggested to used combination of mixed and single database for more accurate DFF prediction.

1. Introduction
Dietary fiber feed (DFF) is a crucial component in ruminant ration [1]. DFF contributed as the main energy source in a ruminant ration, especially grazing ruminants [2]. Dairy cattle used more than 50% DFF in their rations [3]-[6]. DFF used included domesticated grasses such as Napier grass and Brachiaria sp., natural grass, or agricultural by-product [7]. The contribution of natural grass and the agricultural by-product was 16 – 30% in dairy cattle ration in Indonesia, which was higher during the dry season [8]. DFF was also used as protein, vitamin, and mineral [9] sources. The DFF quality used in dairy ration influenced dairy performances [10].

The DFF quality can be measured using chemical contents and fiber fractions [11]. The primary parameters used include proximate compositions such as DM, ash, crude protein (CP), ether extract...
(EE), crude fiber (CF), or plant cell wall structures such as neutral detergent fiber (NDF), acid detergent fiber (ADF), acid detergent lignin (ADL) and silica to calculate hemicellulose, cellulose and lignin fractions of the DFF[7]. The most common analysis of forage quality is wet chemical proximate [12] and Van Soest [13] analysis. The methods are destructive to samples, produce chemical waste, require skilled personnel, less practical, expensive, and require a substantial of time [14].

Near Infrared Reflectance Spectroscopy (NIRS) is an alternative method in predicting the nutritional value of DFF. NIRS is fast, non-destructive, does not require chemicals and leaves chemical waste, and is more cost-effective [15]. NIRS has been widely used in studies related to forage to estimate the chemical composition of forages [16], predict digestibility [17], and milk fatty acid compositions [18]. The accuracy of the NIRS method depends on the database used in the calibration process by the NIRS instrument [19].

Development of a local database that was more relevant to the analyzed feedstuffs improved prediction accuracy of DFF nutritive value [7]. However, the developed databases show low accuracy in the detection of complex organic substances such as ADF and NDF. This result is presumably because of mixed DFF species used in the database development. Therefore, this study aimed to compare the accuracy of single species and mixed-species NIRS databases in predicting the nutrient content of local DFF used in dairy farming in Indonesia.

2. Materials and methods

2.1 Sample preparations

Five main DFFs used in Indonesian dairy farms were used in the development of a mixed-species database which included Napier grass, rice straw, corn stover, natural grass, and cornhusk. The DFF were collected from Pangalengan District of Bandung Regency, Lembang District of West Bandung Regency, Parung Kuda District of Sukabumi Regency, and Cibungbulang District of Bogor Regency. All sampling areas were located in the West Java Province of Indonesia. For each area, two samples of each DFF were collected. Thirty samples of Napier grass from thirty different locations were collected for single-species database development. For external validation, seven completely independent samples of Napier grass were used. About 2 kg of a fresh sample of each were collected and processed in the laboratory.

2.2 Wet chemical analysis

Fresh DFF samples taken from the dairy farm were chopped into 2 cm length, dried in Eyela NDO 400 (made in Japan) oven (at 60°C for 48 hours), ground using laboratory blender, and then filtered to pass a 1 mm screen. Proximate analyses to measure DM, Ash, CP, and CF were conducted according to the AOAC procedure [12]. The DMs were determined after samples were dried in Eyela NDO 400 (made in Japan) oven at 105°C for 48 hours. The CPs were determined using Kjeldahl systems from Gerhart Instruments (made in Germany).

Ankom Fiber Analyzer A200 (made in the US) was used to determine crude fiber and cell wall structure. The amount of 0.5 g of each sample was placed in predetermined weight of F57 Ankom filter bag. The bags were sealed using Impulse sealer AIE 200-2 (made in the US) and distributed evenly into Ankom suspender tray. The tray was then set into an Ankom digestion instrument. The amount of 2 L of NDS was poured into the digestion mechanism followed 4 ml alpha-amyrase and 20 g sodium sulfite. The digestion instrument was closed and then the program is run (temperature 100°C, 75 minutes). Once the program has stopped, the NDS solution was released and the bags were washed thrice with 70 – 90°C hot water and solved in acetone to remove water residue. The bags were then put into an aluminum tray. Before inserting the tray into the oven, it should be assured that all acetone had been evaporated in a fume hood. The bags were then dried for 4 hours and weighted. ADF and crude fiber determinations followed a similar procedure as NDF. However, the solutions used were different. ADF determination used ADS solution, while crude fiber used acid and alkaline solutions. The CFs were determined using method Ba 6a-05 approved by AOCS [20]. The NDF and
ADF were measured according to method 15 and method 14 respectively (Ankom200 manual). Calculation of NDF and ADF fraction was done according to Van Soest [13].

2.3 Measurement feed quality using NIRS
The DFF database was developed using FT-NIR Spectrometer Solids Cell (BUCHI; NIRFlex N-500 made in Switzerland). Before collecting the sample spectrum, NIRS performance was adjusted by running an automatic system suitability test (SST) followed by internal reference measurement at NIRS ware operator application. The external reference tests were conducted by inserting an external reference (provided by BUCHI) into the external-reference holder. Once the reference test finishes, 50 g dried and mashed DFF is distributed evenly into a 100 mm diameter petri dish and then inserted into the petri dish holder. The near-infrared light at 800–2500 nm or 12500–4000 cm⁻¹ was sent into the sample for scanning. Sample identification was carried out by measuring NIRS absorbance after penetrating it up to several mm deep. Each sample went through the scanning process three times. The collected spectra were input with chemo-metric results at NIRSware Management Console application. The NIRCal V5.6 application was used for calibration and validation of the database. Using block-wise method, the collected spectra were divided into calibration data (2/3) and validation data (1/3). Calibration models were developed using partial least square (PLS) regression, while validation models were developed using a validation set.

Statistic parameters were compared between single species and mixed-species database developed models. According to the models generated, the best model was chosen for minimum standard error of calibration (SEC) and standard error of prediction (SEP) with maximum calibration coefficients (R2) and residual predictive deviation (RPD). The RPD was calculated as a ratio between the standard deviation (SD) to SEP. External validation was conducted by measuring samples using a new database and comparing them with chemometrics results. The comparison was made by calculating the ratio between the standard error of prediction to the standard error of laboratory (SEP/SEL).

3. Results and Discussion
3.1 Sample spectrum
The spectrum generated from the single species and mixed-species is shown in figure 1 and figure 2. The spectrum is in the wavelength range of 4000 to 10000 cm⁻¹. The figure shows the significant difference between peaks and valleys of absorption of single species in comparison to mixed species. According to Yang et al., [21], the difference in the height and low of the spectra results showed the differences in chemical components of the samples because different chemical bonds absorbed different wavelengths [22]. In general, all the spectra samples tested had a similar graphic pattern, indicating that the NIRS instrument was able to provide consistent measurement results. According to Stuth et al., [15], the shape of spectra was influenced by the instrument used, sample particle size, chemical concentration in the sample, and wavelength used. With a similar instrument, sample particle size, and wavelength used, it is assumed that the differences in the spectra shape were caused by the different concentrations of chemicals in the sample. As we can see from the figure, single-species spectra were more different in comparison to mixed species. The broad spectra in single-species might be the result of the difference in the variety of Napier grass used. Napier grass was reported to have a significantly different yield and quality between varieties and cultivars [23], [24].
3.2 Calibration and validation

The calibration and validation of mix and single species databases are shown in Table 1. In general, the table shows that statistic parameters of the mixed-species model were better in comparison to the single species model. The R²C in the mixed-species calibration was higher in all parameters observed, except for CF. Having both databases in NIRS can improve DFF nutrient value determination. The highest R²C (0.913) was found in NDF determination using the mixed-species model. This value was higher in comparison to R²C for NDF (0.891) using a mixed model reported by Despal et al. [7]. This might be the result of the less different mixed-species sample used.

Table 1. Calibration and validation statistics of mixed-species and single species for estimation of nutrient content (DM, Ash, CP, CF, ADF, and ADF).

| Parameters | Calibration | Validation |
|------------|-------------|------------|
|            | n  | Mean | Range | SD | SEC | R²C | RPD | n  | Mean | Range | SD | SEP | R²V | RPD |
| Single Species (Napier Grass) | | | | | | | | | | | | | | |
| DM         | 60 | 90.59 | 85.64-94.62 | 2.080 | 1.290 | 0.615 | 1.612 | 30 | 90.20 | 85.64-94.62 | 2.474 | 1.410 | 0.566 | 1.755 |
| Ash        | 60 | 13.63 | 9.92-24.49 | 3.857 | 1.699 | 0.806 | 1.703 | 30 | 14.35 | 10.27-19.49 | 2.894 | 1.623 | 0.668 | 1.783 |
| CP         | 60 | 10.31 | 4.71-14.61 | 2.583 | 1.042 | 0.837 | 2.479 | 30 | 11.11 | 9.95-14.02 | 2.583 | 1.141 | 0.564 | 2.264 |
| CF         | 60 | 28.20 | 17.3-35.8 | 5.064 | 1.919 | 0.856 | 2.639 | 30 | 28.81 | 20.49-32.1 | 3.065 | 1.882 | 0.646 | 1.629 |
| NDF        | 60 | 58.21 | 33.8-65.58 | 7.687 | 2.495 | 0.895 | 3.081 | 30 | 59.32 | 51.56-62.56 | 2.661 | 2.278 | 0.529 | 1.168 |
| ADF        | 60 | 33.24 | 18.63-26.17 | 5.111 | 2.592 | 0.670 | 1.972 | 30 | 34.01 | 26.17-39.67 | 3.040 | 2.274 | 0.670 | 1.337 |

Mixed-species

| Parameters | Calibration | Validation |
|------------|-------------|------------|
|            | n  | Mean | Range | SD | SEC | R²C | RPD | n  | Mean | Range | SD | SEP | R²V | RPD |
| DM         | 60 | 90.63 | 87.35-93.15 | 1.256 | 0.618 | 0.758 | 2.032 | 30 | 89.80 | 85.17-91.93 | 1.988 | 0.952 | 0.779 | 2.087 |
| Ash        | 60 | 12.39 | 3.94-20.65 | 4.874 | 1.788 | 0.865 | 2.725 | 30 | 8.66 | 2.43-19.66 | 5.663 | 1.928 | 0.888 | 2.937 |
| CP         | 60 | 8.15  | 4.74-11.55 | 2.218 | 0.860 | 0.860 | 2.580 | 30 | 8.33 | 5.91-11.4 | 2.021 | 0.766 | 0.850 | 2.639 |
| CF         | 60 | 29.37 | 24.35-33.5 | 2.988 | 1.460 | 0.761 | 2.046 | 30 | 28.82 | 24.8-32.2 | 2.713 | 1.447 | 0.722 | 1.874 |
| NDF        | 60 | 61.59 | 51.81-77.41 | 6.380 | 1.887 | 0.913 | 3.380 | 30 | 62.22 | 54.7-72.1 | 5.823 | 1.853 | 0.931 | 3.143 |
| ADF        | 60 | 34.46 | 27.73-41.11 | 4.003 | 1.475 | 0.864 | 2.713 | 30 | 32.97 | 27.87-38.04 | 3.053 | 1.470 | 0.812 | 2.077 |

Note: n= total number of observation, DM= dry matter, CP= crude protein, CF= crude fiber, NDF= neutral detergent fiber, ADF= acid detergent fiber, SD= standard deviation, SEC= standard error of calibration, R2C= coefficient determination of calibration, RPD= residual predictive deviation, SEP= standard error of prediction, R2V= coefficient determination of validation, * = p<0.05.

Calibration is a process that relates the chemical information contained in the spectral properties of a substance with chemical (or physical) information obtained from laboratory analysis. The purpose of the calibration is so that the user can calculate the constituents of interest using only NIRS, without...
carrying out laboratory analysis. The calibrations should be well distributed, representing possible variations. These variations can be (1) temporal, for example the time of collection; (2) spatial, for example the location range or geographic location; or (3) biological, for example cultivar or growth stage [15]. However, a wide different variation of samples might result in a varied concentration of chemical content and broad-spectrum [21], lower R²C that resulted in less accurate prediction [25]. In general, validation of the database resulted in better R²V for mixed species but lower R²V for single species.

The standard error of calibration (SEC) of the single species model was higher than the mixed model for all parameters observed, except for CP. The higher SEC value represented the more significant difference between laboratory and NIRS prediction results [26]. This result showed that mixed-species models were more accurate than single species models. The standard error of predictions (SEP) was also lower in mixed-species models compared to the single species models except for parameter ash that was higher in the mixed model. SEP values represent the accuracy of the validated model [27].

The RPD value in the table shows that mixed-species model produced RPD > 2 for all parameters, while, in the single species model, RPD > 2 were found in CP, CF, and NDF parameters only. According to [28], RPD represents the ability of the NIRS model to predict a substance. In this case, mixed-species model predicted DFF value more accurately. Moreover, [28] stated that RPD values in the range 1.5-1.9 indicate rough predictions so that calibration is needed, range 2-2.5 showed a decent predictive model, and values above 2.5 indicate an excellent predictive model. In the mixed-species calibration set, RPD on ash, CP, NDF, and ADF were above 2.5. This result shows that the NIRS mixed-species model was an excellent good predictor for ash, CP, NDF, and ADF nutrient contents in DFF.

Table 2. External validation statistics of nutrient contents on mixed-species and single species.

| Parameters | Single Species | Mixed-species |
|------------|----------------|---------------|
|            | WCA | NIRS | | WCA | NIRS | | | | | | |
|            | AVG  | SD   | AVG  | SD   | AVG  | SD   | AVG  | SD   | T-Test | R²   | SEL | SEP | SEP/SEL |
| DM         | 91.246 | 2.386 | 90.525 | 2.408 | 0.583 | 0.338 | 1.642 | 1.914 | 1.166 |
| Ash        | 14.604 | 2.001 | 14.246 | 1.882 | 0.737 | 0.914 | 1.130 | 0.640 | 0.566 |
| Protein    | 10.991 | 1.215 | 9.299 | 1.746 | 0.065 | 0.044 | 0.892 | 2.056 | 2.306 |
| CF         | 29.414 | 3.014 | 29.385 | 2.773 | 0.985 | 0.494 | 0.918 | 2.173 | 2.366 |
| NDF        | 34.614 | 2.871 | 32.990 | 3.596 | 0.108 | 0.344 | 1.365 | 2.772 | 2.031 |
| ADF        | 61.171 | 2.939 | 58.466 | 2.890 | 0.369 | 0.552 | 1.047 | 2.452 | 2.342 |

External validation using Napier grass samples from different locations is presented in Table 2. The table shows that the SEP/SEL < 2 were found in all parameters using the mixed-species database,
except for CF and ADF. While the single species model produced SEP/SEL > 2 in all parameters except for DM and ash. The smaller SEP/SEL the more accurate the model. Assessment predictive ability of selected calibration equation should be carried out using an independent sample test (external validation). The samples need to be different from the calibration set [29, 30]. External validation was a better tool to test the performance of the prediction models [16].

Overall, from the observed parameters, it appears that the NIRS database for the mixed-species was better at predicting the nutritional value of DFF compared to single species. This finding cannot approve the hypothesis. This might be caused by a large variety of nutrient content in Napier grass due to different varieties and cultivar used. The small number of samples used might also contribute to the lower accuracy of single species models. Although, the BUCHI manual required 30 samples for minimum calibration, the more sample used, the more change to fulfill calibration sample requirements such as being well distributed and representing possible variations [29].

4. Conclusion
A combination of mixed-species and single species NIRS databases can be used in predicting the nutrient value of dairy fiber feed for daily use. The used combination of both databases improved the prediction accuracy of dietary fiber feed for dairy. Single species feed alone cannot improve prediction accuracy in all parameters. It is suggested to use both databases in dietary fiber feed analysis using NIRS.

References
[1] Despal, Permana I G, Toharmat T and Evvienrie DEA 2017 Pemberian Pakan Sapi Perah 1st ed. (Bogor: IPB Press)
[2] Manousidis T, Parissi Z M, Kyriazopoulos A P, Malesios C, Koutroubas S D and Abas Z 2018 Livest. Sci. 218, 8–19
[3] Lestari D A, Abdullah L and Despal 2015 Media Peternakan. 38, 110-117
[4] Nugroho H D, Permana I G and Despal 2015 Media Peternakan. 38, 70-76
[5] Zahera R, Permana I G and Despal 2015 Media Peternakan. 38, 123-31
[6] Hasanah U, Permana I G and Despal 2017 Pakistan J. Nutr. 16, 577-587
[7] Despal, Sari L A, Chandra R, Zahera R, Permana I G and Abdullah L 2020 Trop. Anim. Sci. J. 43, 263–9
[8] Despal, Malyadi J, Destianingsih Y, Lestari A, Hartono H and Abdullah L 2014 Proc. of the Int. Workshop Tropical Bio-Resources for Sustainable Development, Bogor, 145-150
[9] Permana I G and Despal 2005 J. Agric. Rural Dev. Trop. Subtrop. Beihf. 88, 37–44
[10] Monzón-Gil E, Castañón J I R and Ventura M R 2010 Small Rumin. Res. 94, 196–200
[11] Fulkerson W J, Neal J S, Clark C F, Horadagoda A, Nandra K S and Barchia I 2007 Livest. Sci. 107, 253–64
[12] Association of Official Analytical Chemist 2015 Official Methods of Analysis of AOAC International 20th ed (Arlington, USA: Assoc. Off. Anal. Chem)
[13] Van Soest P J, Robertson J B and Lewis B A 1991 J Dairy Sci. 74, 3583–97
[14] Parrini S, Acciaioli A, Crovetti A and Bozzi R 2017 J. Anim. Sci. 17, 87–91
[15] Stuth J, Jama A and Tolleson D 2003 F. Crop Res. 84, 45–56
[16] Lobos I, Gou P, Hube S, Saldaña R and Alfar M 2013 J Soil Sci. Plant. Nutr. 13, 463–468
[17] Samadi, Wajizah S and Munawar A A 2018 Trop. Anim. Sci. J. 41, 121–127
[18] González-Martin M I, Vivar-Quintana A M, Revilla I and Salvador-Esteban J 2020 Microchem. J. 156, 104854
[19] Hall M B 2014 Vet. Clin North Am - Food. Anim. Pract. 30, 487–505
[20] Association of Official Analytical Chemist 2005 Official Methods and Recommendation Practices of the AOCS 7th ed. (Urbana, USA: The American Oil Chemists’ Society)
[21] Yang Z, Nie G, Pan L, Zang Y, Huang L, Ma X and Zhang X 2017 Peer. J. 5:e3867
[22] Ozaki Y, Morita S and Du Y 2007 Spectral analysis. In: Ozaki Y, McClure W and Christy A (eds) Near-Infrared Spectroscopy in Food Science and Technology (New Jersey, AS: John Wiley & Sons Inc.)

[23] Zailan M Z, Yaakub H and Jusoh S 2018 J. Anim. Plant. Sci. 28, 63–72

[24] Antony S and Thomas CG 2014 J. Trop. Agric. 52, 90–3

[25] Yin Y 2020 J. Stat. Plan. Inference. 206, 79–195

[26] Liu X, Han L, Yang Z and Xu Ch 2008 J. Anim. Feed. Sci. 17, 631–9

[27] Williams P C and Sobering D C 1993 J. Near Infrared Spectrosc. 1, 25–32

[28] Mouazen A M, Saëys W, Xing J, De Baerdemaeker J and Ramon H 2005 J. Near Infrared Spectrosc. 13, 87–97

[29] Stuth J, Jama A and Tolleson D 2003 Field Crops Research 84, 45–56

[30] Munawar AA, von Hörsten D, Wegener JK, Pawelzik E, Mörlein D 2016. Eng Agric Environ Food. 9(3):208–215.

Acknowledgment
This research was funded by Indonesian Ministry of Education within the Scheme of National Competitive Applied Research 2020 with contract No 3984/IT3.L1/ PN2020.