Revista Virtual de Química

ISSN 1984-6835

Artigo

Chemometric Determination of Cellulase and Xylanase Enzyme Activities Based on Near Infrared Spectroscopy: A New Approach

Chaves, A. R.;* Pinheiro, T. S.; Pantoja, L. A.; dos Santos, A. S.

Rev. Virtual Quim., **2021**, *13* (2), *no prelo*. Data de publicação na Web: 8 de Fevereiro de 2021

http://rvq.sbq.org.br

Determinação Quimiométrica de Atividades Enzimáticas de Celulases e Xilanase Utilizando Espectroscopia de Infravermelho Próximo: Uma Nova Abordagem

Resumo: O uso da espectroscopia na região do infravermelho próximo, combinada com metodologias de análise multivariada, tem permitido a análise quantitativa, direta e não destrutiva de várias substâncias orgânicas e inorgânicas em amostras complexas. Neste trabalho foi investigado o possível uso da abordagem quimiométrica para a determinação direta das atividades enzimáticas de celulases e xilanase com base na correlação matemática entre dados obtidos por metodologia tradicional e os obtidos por espectroscopia de infravermelho próximo. Para este fim, um produto comercial contendo as enzimas alvo foi tratado termicamente para a construção de variabilidade amostral. As atividades de FPase, CMCase, β-glucosidase e xilanase foram determinadas por métodos colorimétricos convencionais. A regressão parcial por mínimos quadrados foi adotada como ferramenta matemática para o estudo de correlação de dados. O coeficiente de determinação obtido para a validação das regressões realizadas variou de 0,84 a 0,92. Todos os valores da relação desempenho / desvio foram > 2,5. A seleção das bandas espectrais foi essencial para o sucesso dos ajustes das regressões. A possibilidade de determinação direta de atividades enzimáticas em solução, aqui retratada, representa uma prova de conceito que visa dispensar o processo de catálise como etapa necessária para a determinação analítica das atividades enzimáticas.

Palavras-chave: Atividade enzimática; análise multivariada; quimiometria; espectroscopia de infravermelho.

Abstract

The use of spectroscopy in the near infrared region, in combination with multivariate analysis methodologies, has enabled the quantitative, direct and non-destructive analysis of several inorganic and organic substances in complex samples. The aim of the current study was to investigate the possible use of chemometric approach to directly determining xylanase and cellulase activities based on mathematical correlation between data from traditional methods and the ones from direct near infrared spectroscopy of enzyme samples. Commercial product comprising the target enzymes was subjected to heat treatment for construction of sample variability. FPase, CMCase, β-glucosidase and xylanase activities were determined by conventional colorimetric methods. Partial least squares regression was adopted as mathematical tool to correlate the analyses. The determination coefficient recorded to validation of regressions ranged from 0.84 to 0.92; all ratio of performance to deviation values were > 2.5. Spectral band selection was essential to enable successful regression adjustments. The likelihood of directly determining enzymatic activity in solution represents an innovative concept, since it does not require catalysis process as necessary step to analytically determine the enzymatic activity.

Keywords: Enzymatic activity; multivariate analysis; chemometrics; infrared spectroscopy.

* Universidade Federal dos Vales do Jequitinhonha e Mucuri, Instituto de Ciência e Tecnologia, CEP 39100-000, Diamantina–MG, Brasil.

amanda.chaves@ufvjm.edu.br

DOI: 10.21577/1984-6835.20210005

Volume 13, Número 2 Março-Abril 2021

Revista Virtual de Química
ISSN 1984-6835

Chemometric Determination of Cellulase and Xylanase Enzyme Activity Based on Near Infrared Spectroscopy: A New Approach

Amanda Rocha Chaves,[a]* Thais Santos Pinheiro,[a] Lílian de Araújo Pantoja,[b] Alexandre Soares dos Santos[c]

[a] Federal University of Jequitinhonha and Mucuri Valleys, Graduate Program in Biofuels, CEP 39100-000, Diamantina–MG, Brazil.

[b] Federal University of Jequitinhonha and Mucuri Valleys, Institute of Science and Technology, CEP 39100-000, Diamantina–MG, Brazil.

[c] Federal University of Jequitinhonha and Mucuri Valleys, Department of Basics Science, CEP 39100-000, Diamantina–MG, Brazil.

*amanda.chaves@ufvjm.edu.br

Recebido em 30 de Junho de 2020. Aceito para publicação em 21 de Dezembro de 2020.

1. Introduction

The activity of enzymes such as cellulases and hemicellulases is traditionally determined through the colorimetric quantification of products released under predefined time and reactional conditions, based on univariate measurement principles.[1-4] Enzymes, as molecular entities, are not the target of direct examination in this analytical approach. Overall, traditional methodologies are time-consuming and expensive, mainly with regards to reagents, equipment and human resources used in the analytical process.[5] A new fast, cheap and robust methodology based on multivariate analysis of data associated with instrumental spectroscopy methods has enabled advancements in biological sample analysis.[6-8] This new analytical approach is called Chemometrics; it is a discipline based on the use of a set of mathematical, statistical and computational tools used to develop multivariate strategies for chemical data evaluation with qualitative and quantitative applications.[9] The analytical methods based on chemometrics provides to reduce time, reagent consumption and number of steps in the analysis of complex samples based on analytical process automation.[10,11] In theory, any analytical technology, such as infrared spectroscopy, fluorescence spectroscopy, nuclear magnetic resonance spectroscopy, mass spectrometry,

UV-Vis spectrophotometry and chromatography, can be used in combination with chemometrics in order to predict analytes' physical, chemical or biological features.[12] There are records in the scientific literature about the use of chemometrics to determine enzymatic activity;[13-15] however, these initiatives target the substrate or products of enzymatic catalysis as analyte, in the same way it is done in traditional methodologies. This chemometrics-use way disregards the chemical nature of enzymes and their ability to respond to the same chemometric methods used to quantify the product of enzymatic catalysis. Infrared spectroscopy is one of the most used techniques in chemometric models adopted for quantitative determination purposes. The infrared spectrum analysis of proteins generate data with high information content. This factor turns infrared spectroscopy into a valuable tool to investigate protein structure.[16] Infrared spectroscopy also has the advantage of being a fast and non-destructive method that does not require reagents and allows simultaneously detecting several analytes.[11] The aim of the current study was to investigate the possibility of directly and simultaneously determining the enzymatic activity of cellulases and xylanases from commercial enzymatic preparation, based on near infrared spectroscopy and chemometrics principles, as well as to present a new approach to the enzyme dosage-development method.

2. Materials and Methods

2.1. Enzymes and heat treatment

Commercial enzymatic cocktail Celluclast® 1.5 L (Novozymes™, Denmark) was used as source of cellulase and xylanase enzymes evaluated in the current study. The aforementioned commercial enzyme preparation was subjected to heat treatment (at 70ºC) in thermostatic water bath in order to generate the variability in the target enzyme activity values. Aliquots of 2 mL of heat-treated samples were recovered on centrifuge plastic microtubes, cooled in ice bath for 30 seconds and centrifuged at 10,000 rpm for 10 minutes, at three-minute intervals. The recovered supernatants were ice stored for up 60 minutes for subsequent analytical determinations.

2.2 Enzyme activity determinations

The activity of FPase, CMCase (endoglucanase), β-glucosidase and xylanase were determined based on the methodology originally suggested by IUPAC [1] and Ghose and Bisaria,[17] which was herein adapted to 96-well microplate scale, with minimal modifications. All assays performed for enzymatic activity determination comprised seven repetitions. The blank test was performed by replacing the enzymes with distilled water. Each assay conducted on microplate was followed by analytical curves using glucose and xylose (0 to 1 g L^{-1}) as standard.

2.2.1. FPase

FPase activity was determined by using Whatman N. 1 filter paper strips (1.0 x 0.3 cm) deposited on the walls of microplate wells. 50 µL of 50 mmol L^{-1} sodium acetate buffer, at pH 4.8, and 50 µL of enzyme solution were added to each well. The enzymatic hydrolysis of the filter paper was maintained under heating in oven at 50ºC for 60 minutes. Next, 100 µL of 3,5-dinitrosalicylic acid reagent (DNS) was added to each well to determine the reducing sugars.[18] Subsequently, the microplate was incubated in water bath at 70ºC for 30 minutes. Finally, it was cooled in ice bath and read in microplate reader (Biochrom, model Asys UVM 340) at 540 nm. One FPase Unit was defined as the amount of enzyme capable of releasing 1 µmol of glucose-equivalent per minute under assay conditions.

2.2.2. CMCase

CMCase activity was determined through the addition of 50 µL of 1% carboxymethylcellulose solution (50 mmol L^{-1} acetate buffer, pH 4.8) to the microplate wells, which were subsequently added with 50 µL of enzyme solution. Oven incubation was carried out at 50ºC for 30 minutes to enable enzymatic hydrolysis of CMC; next, 100 µL of DNS reagent were added to each well. The microplate was incubated in water bath at 70ºC for 30 minutes, cooled in ice bath and read in microplate reader (Biochrom, model Asys UVM 340) at 540 nm. One CMCase Unit was defined as the amount of enzyme capable of releasing 1 µmol of glucose-equivalent per minute under assay conditions.

2.2.3. β-glucosidase

β-glucosidase activity was determined by mixing 50 µL of 15 mmol L^{-1} cellobiose solution (50 mmol L^{-1} sodium acetate buffer, pH 4.8) and 50 µL of enzyme solution in microplate wells. The plate was subjected to incubation at 50ºC for 10 minutes, transferred to water bath with boiling water for 1 minute and cooled in ice bath. The released glucose was quantified based on the GOD-POD enzymatic-colorimetric method by Lloyd.[19] One unit of β-glucosidase was defined as the amount of enzyme capable of releasing 1 µmol of glucose per minute under assay conditions.

2.2.4. Xylanase

Firstly, 50 µL of birchwood xylan solution (1.4% in 50 mmol L^{-1} sodium acetate buffer, pH 4.8) was mixed with 50 µL of enzyme solution in microplate wells. Microplates were incubated in oven at 50ºC for 5 minutes to enable the enzymatic hydrolysis of birchwood xylan. Next, 100 µL of DNS reagent were added to each well; reducing sugars' reaction was performed in water bath at 70ºC for 30 minutes. Microplate was left to cool in ice bath and read in microplate reader (Biochrom, model Asys UVM 340) at 540 nm. One xylanase Unit was defined as the amount of enzyme capable of releasing 1 µmol of xylose-equivalent per minute, under assay conditions.

2.3. Near infrared spectroscopy

The very same eleven samples deriving from the pretreated enzymatic preparation used for traditional enzymatic activity determinations, as described in 2.1 item, were analyzed in near infrared spectrophotometer (model NIR 9000 PLS, FEMTO, Brazil) after 100X diluted with distilled water. The parameters of the acquisition were spectral width from 1100 to 2500 nm, transmittance mode at 2-nm resolution and it generated 700 absorbance values within 70 seconds. Three spectra were generated for each sample, totalizing 33 spectra.

2.4. Spectral data processing and multivariate analysis

Unscrambler X software (Camo Software, Oslo, Norway) was used for data processing and multivariate analyses. All primary spectral data were pre-processed based on the Savitzky-Golay smoothing,[20] Standard Normal Variate (SNV) transformation and detrending methods.[21] Partial Least Square Regression (PLS) was the algorithm used for multivariate data analysis.[22] The y and x components of data matrix were represented by enzyme activities of all samples and by all NIR spectra, respectively. To validate the model performance was used a cross-validation approach to split the data into training and test sets for simulated and real data set with 20 segments choose randomly. Outliers were identified and removed through the combined observations of the samples behavior on plot of Y residual values versus predicted Y values, normal quantile plot (Q-Q plot) and score plot (Hoetlling T2 ellipse) constructed according to the modeling of each enzyme activity generated with optimal number of factors at 5% significance level. The quality of PLS regression-based models was evaluated through RMSEC (Root Mean Square Error of Calibration), RMSECV (Root Mean Square Error of Cross Validation), RPD (Ratio of Performance to Deviation) and RER (Range Error Ratio) analysis. RPD was calculated as the ratio between the standard deviation of the reference data for the validation set and the standard error of prediction (from cross-validation). RER was calculated as the ratio between the range in validation reference data and the standard error of prediction (from cross-validation).[23]

3. Results and Discussion

Enzymatic activity values recorded for FPase, CMCase, xylanase and β-glucosidase found in the commercial enzymatic preparation Celluclast® are shown in table 1. Collected data allowed observing that the enzymatic cocktail presented the greatest relative amount of xylanase activity (1.043 U mL^{-1}); it was followed by β-glucosidase (254 U mL^{-1}), CMCase (155 U mL^{-1}) and FPase (68 U mL^{-1}). Coefficient of variation rates (CV%) recorded in the colorimetric analytical determinations adapted to the microplate scale ranged from 2.1 to 5.9. The observed variance was comparable to those observed in other methodological proposals;[24-27] it was considered acceptable for enzyme activity determinations.

Table 1. Mean FPase, CMCase, β-glucosidase and xylanase enzymatic activity values recorded for enzyme preparation Celluclast®

Enzymatic Activity (U mL^{-1})				
Values	**β-glucosidase**	**CMCase**	**FPase**	**Xylanase**
Min	248.61	144.64	67.13	1018.64
Max	273.58	162.12	69.85	1060.04
Average	257.74	155.06	68.88	1043.77
CV%	5.34	5.94	2.20	2.12

U: amount of enzyme capable of forming 1 µmol of the evaluated product per minute under assay conditions; CV: Coefficient of variation

Heat treatment application to commercial preparation Celluclast® at 70 °C for up to 30 minutes resulted in 11 samples, whose target enzymatic activity was determined based on the traditional approach, as previously described in section 'Methodology'. Gradual thermal denaturation process enabled obtaining samples with the following minimum and maximum enzymatic activity and coefficients of variation: FPase (14.2 to 81.3 U mL^{-1}, CV=0.67), CMCase (97.1 to 186.1 U mL^{-1}, CV=0.25), β-glucosidase (83.9 to 285.9 U mL^{-1}, CV=0.32) and xylanase (144.9 to 1,088.5 U mL^{-1}, CV=0.71). Another notable phenomenon observed in the thermal denaturation process lies on the selective way, according to which, the enzymatic activity decay took place (Figure 1). Xylanolytic activity was mostly affected by the heat treatment in the first half of the exposure time, whereas CMCase activity was the least affected throughout the treatment period. On the other hand, β-glucosidase activity was mostly affected in the second half of the heat exposure time. This asymmetric behavior has generated samples with different enzyme activity rates. These results were essential to generate sample variability, which allowed maintaining the fortuitous nature of the enzyme solution matrix. If this process had happened in a homogeneous way, as it happens in dilution processes, it would masking the contribution of target analytes.

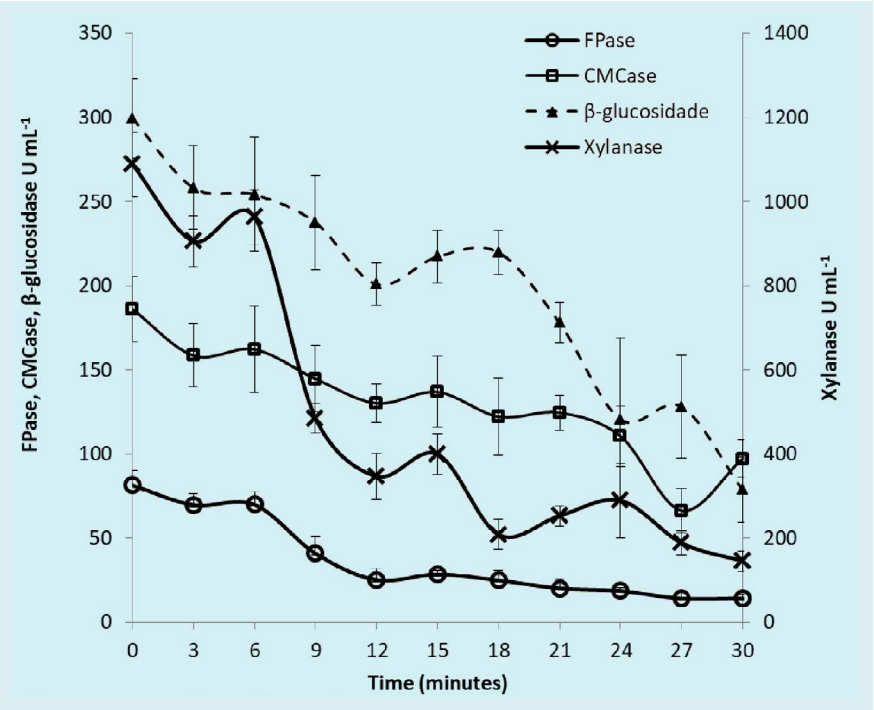

Figure 1. Heat treatment progress and its effect on enzyme activity. Vertical bars represent the standard deviation to each activity determination

The very same samples obtained through the heat treatment have produced 33 infrared spectra in the 1500 to 2500 nm horizon, when they were subjected to near infrared spectroscopy, including the repetitions performed in triplicate. The spectroscopic profile of the samples is shown in figure 2A. It was possible visually highlighting two spectral areas presenting large irregular signal variation in regions ranging from 1862 to 2032 nm and from 2288 to 2500 nm. These two regions have shown spectral behavior typical of noise and signal dispersion. Non-linear Iterative Partial Least Square algorithm (NIPALS) application to investigate correlation between independent variables (response), represented by the enzymatic activities evaluated (FPase, CMCase, xylanase and β-glucosidase), and the predictors, by taking into consideration all spectral data set comprising all readings from 1500 to 2500 nm, resulted in very low coefficients of determination (R^2) and very high root-mean-square error (RMSE) values (Table 2). The first strategy used to make the set of predictors more appropriate to model the correlation to independent variables lied on removing the noise region. This region became evident due to the application of data transformation by second order derivative (Figure 2B), which, in its turn, led to increased noise-to-signal ratio. Spectral data were scaled to the 1100-1830 nm range. There was significant improvement in values recorded for R^2 and RMSE when the PLS algorithm (NIPALS) was run with the resized data set, although without any further processing (Table 2).

Table 2. Analysis of the calibration and validation of models used to predict enzymatic activity through near infrared spectroscopy based on partial least squares regression

PLS	Measures	Enzymes			
		β-glucosidase	Fpase	Cmcase	Xylanase
Without resizing	Samples	33	33	33	33
	Factors	1	1	1	1
	R^2_{Cal}	0.534	0.489	0.518	0.455
	R^2_{Val}	0.508	0.453	0.463	0.424
	RMSEC	42.19	16.93	21.80	240.52
	RMSEV	47.04	12.37	24.26	266.67
	$Bias_{Val}$	0.42	-0.04	-0.003	-0.27
	RPD	1.31	1.26	1.29	1.22
	RER	4.29	5.42	3.67	3.54
With resizing and Without pre-processing	Samples	33	33	33	33
	Factors	4	4	4	4
	R^2_{Cal}	0.824	0.926	0.847	0.942
	R^2_{Val}	0.783	0.751	0.771	0.729
	RMSEC	25.93	6.42	12.27	78.28
	RMSEV	29.73	12.37	15.83	172.28
	$Bias_{Val}$	-0.91	-0.63	-0.55	-9.43
	RPD	2.07	1.91	1.98	1.89
	RER	6.66	5.35	5.53	5.40
With resizing and With pre-processing	Samples	30	30	25	27
	Factors	4	7	6	6
	R^2_{Cal}	0.935	0.959	0.987	0.963
	R^2_{Val}	0.857	0.871	0.920	0.839
	RMSEC	13.25	3.01	2.34	57.23
	RMSEV	21.46	8.06	6.63	124.47
	$Bias_{Val}$	-0.83	-0.32	0.07	-7.60
	RPD	2.88	2.61	4.72	2.61
	RER	7.04	8.19	11.91	7.45

R^2_{Cal} = coefficient of determination for calibration; R^2_{Val} = coefficient of determination for validation; PLS*: Partial Least Squa*re; RMSEC= root mean square error of calibration; RMSECV= root mean square error of cross validation; RPD= Ratio of Performance to Deviation; RER= Range Error Ratio

In addition to resizing the spectral data, combinations of mathematical data preprocessing and sample outlier removal were tested to improve the calibration and validation of the regression models. Spectral data preprocessing is the most important step before chemometric bi-linear modeling.[28] Phenomena such as baseline shifts among samples, scatter effects, as well as other unspecific and random noises, can reduce the signal-to-noise ratio (SNR), affect spectra resolution and hinder the accuracy and precision of calibration models.[29,30] Smoothing treatments, normalization by the standard deviation of the signals and removal of systematic shifts inherent to the adopted equipment or to the sample matrix were combined in the current study. The combination of second-order polynomial Savitzky-Golay smoothing transformation with 5-point window, which was followed by Standard Normal Variate transformation (SNV), has resulted in the best PLS regression model adjustment for β-glucosidase activity prediction. The SNV combined with fourth-order polynomial detrend transformation was the most suitable combination

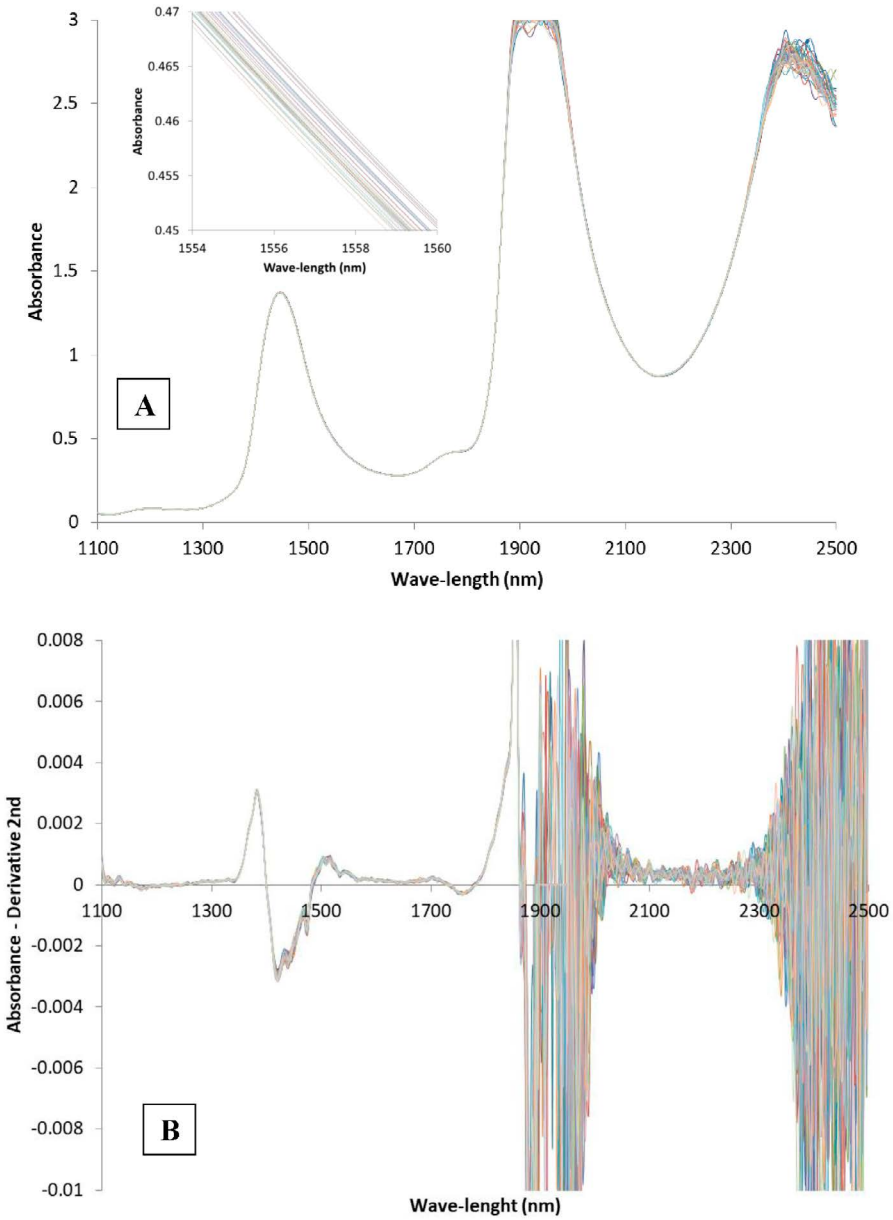

Figure 2. NIR scanning spectra of Celluclast® samples obtained after heat treatment application at 70ºC, for different times. (A). NIR spectra from primary data set. (B). NIR spectra from data transformed by derivative 2nd. The figures present 33 overlapped spectra

to model the PLS regression of FPasic, CMCasic and xylanase activity.

Calibration models used for β-glucosidase, CMCase, FPase and xylanase enzyme prediction presented high coefficient of determination (R^2_{Cal}) (>0.93); this outcome pointed towards the strong interaction between values recorded for enzymatic activities determined through conventional colorimetric methodology and the information available in the NIR spectra region (Table 2). The validation of the regression model applied to the relatively small set of samples evaluated in the current study used 20 segments of randomly crossed samples. Cross-validations have shown coefficient of determination (R^2_{Val}) ranging from 0.84 to 0.92. RMSE values were significantly lower than the ones determined through primary data, without preprocessing.

All regressions were run based on 1 to 7 factors. The increased number of factors (number of latent variables) tended to increase the coefficient of determination. However, the apparent increase in the quality of the regression model also meant increased RMSE values, which is only useful to fit the observations of the learning sample set rather than new observations.[31, 32] The optimal number of factors for each modeling was determined through analysis of residual variance versus RMSE curve generated in the Unscrambler software, observing the break in the decreasing trend of the residual variance. After resizing, predictor preprocessing (NIR data) and outlier removal procedures were over, the number of optimal factors used for modeling regressions increased from 4 to 6 for CMCase and xylanase activity, and from 4 to 7 for FPase activity. However,

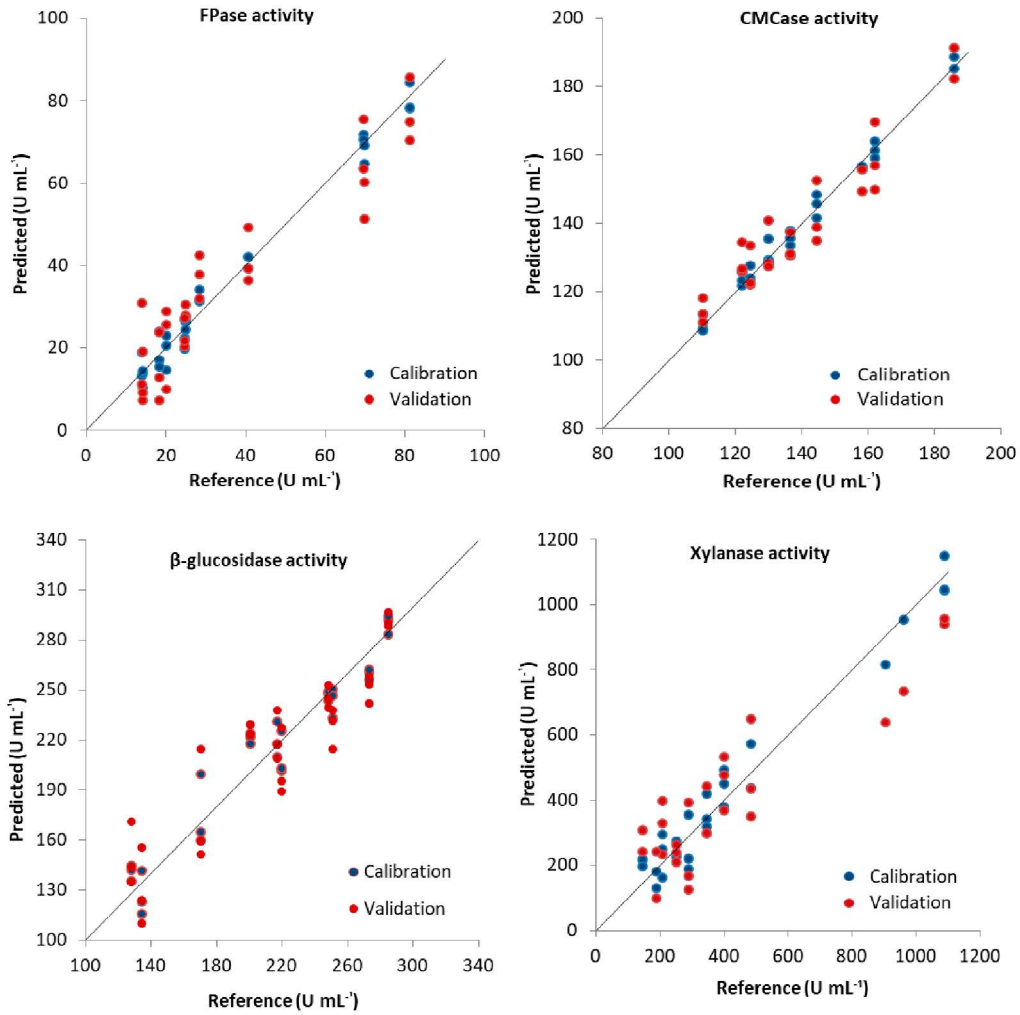

Figure 3. Correlation profile between reference values and values predicted for FPasic, CMCasic, Xylanasic and β-glucosidasic activity through the modeling of NIR spectroscopy data based on PLS

RMSE values decreased, and RPD and RER values significantly increased. This probably happened because was possible maintaining at least four times the number of samples in comparison to the number of factors selected for cross-validation predictions, in all cases. According to Baum *et al.* [11], the number of samples should be at least 3 times larger than the number of latent variables to enable good adjustments in the regression model. All PLS models adjusted for the enzymatic activity predictions evaluated based on cross validation in the present study reached RPD values > 2.6. According to Viscarra Rossel *et al.*,[33] RPD values > 2.5 indicate excellent quantitative model/predictions. In addition, models used to predict FPase and CMCase activity recorded RER values of 8.2 and 11.9, respectively. According to thresholds set by Malley *et al.*,[34] 10 ≤ RER < 15 values are moderately successful, whereas 8 ≤ RER <10 values indicate moderately useful prediction models. Correlation curves between predicted and reference values for regression models generated based on preprocessed spectral data is shown in figure 3. Lines passing through zero are theoretical and illustrate a unitary correlation.

4. Conclusion

Results in the present study enabled concluding that spectroscopy application in the near infrared region to chemometrically quantify the activity of the herein investigated enzymes was capable of producing an acceptable and promising mathematical correlation to develop the original methodology. This approach opens room for the development of direct enzyme activity determination methods based on enzyme observation as molecular entity, rather than on substrates or products of enzyme-catalyzed reactions. The possible adoption of this type of methodological approach implies shorter analysis time, reduced reagent amounts or dispensation, lower generation of chemical waste by analytical processes and likely reduced analysis cost.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Finance Code 001, and by Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

References

[1] Ghose, T. K. Measurement of cellulase activities. *Pure and Applied Chemistry* **1987**, *59*, 257. [CrossRef]
[2] Sharrock, K. R. Cellulase assay methods: a review. *Journal of Biochemical and Biophysical Methods* **1988**, *17*, 81. [CrossRef]
[3] Shuangqi, T.; Zhenyu, W.; Ziluan, F.; Lili, Z.; Jichang, W. Determination methods of cellulase activity. *African Journal of Biotechnology* **2011**, *10*, 7122. [Link]
[4] Dashtban, M.; Maki, M.; Leung, K. T.; Mao, C.; Qin, W. Cellulase activities in biomass conversion: measurement methods and comparison,*Critical Reviews in Biotechnology* **2010**, *30*, 302.[CrossRef]
[5] Siqueira, J. G. W.; Teixeira, N. A.; Vandenberghe, L. P. S.; Oliveira, P. Z.; Soccol, C. R.; Rodrigues, C. Update and Revalidation of Ghose's Cellulase Assay Methodology. *Applied Biochemistry and Biotechnology* **2020**, *191*, 1271. [CrossRef]
[6] Rathore, A. S.; Bhushan, N.; Hadpe, S. Chemometrics applications in biotech processes: A review. *Biotechnology Progress* **2011**, *27*, 307. [CrossRef]
[7] Santos, C. A. T.; Lopo, M.; Ricardo, N. M. J.; Páscoa, J. A. L. A Review on the Applications of Portable Near-Infrared Spectrometers in the Agro-Food Industry. *Applied Spectroscopy* **2013**, *67*, 1215. [CrossRef]
[8] Truong, V. K.; Dupont, M.; Elbourne, A.; Gangadoo, S.; Rajapaksha Pathirannahalage, P.; Cheeseman, S.; Chapman, J.; Cozzolino, D. From Academia to Reality Check: A Theoretical Framework on the Use of Chemometric in Food Sciences. *Foods* **2019**, *8*, 164. [CrossRef]
[9] Brereton, R. G.; Jansen, J.; Lopes, J.; Marini, F.; Pomerantsev, A.; Rodionova, O.; Roger, J. M.; Walczak, B.; Tauler, R. Chemometrics in analytical chemistry—part II: modeling, validation, and applications. *Analytical and Bioanalytical Chemistry* **2018**, *410*, 6691. [CrossRef]
[10] Santos, M. C.; Nascimento, P. A. M.; Guedes, W. N.; Pereira-Filho, E. R.; Filletti, E. R.; Pereira, F. M. V. Chemometrics in analytical chemistry – an overview of applications from 2014 to 2018. *Eclética Química Journal* **2019**, *44*, 11. [CrosRef]
[11] Pasquini, C. Near Infrared Spectroscopy: fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society* **2003,** *14*, 198.[CrossRef]

[12] Singh, I.; Juneja, P.; Kaur, B.; Kumar, P. Pharmaceutical Applications of Chemometric Techniques. *ISRN Analytical Chemistry* 2013, 795178. [CrossRef]

[13] Baum, A.; Hansen, P. W.; Meyer, A. S.; Mikkelsen, J. D. Simultaneous measurement of two enzyme activities using infrared spectroscopy: A comparative evaluation of PARAFAC, TUCKER and N-PLS modeling, *Analytica Chimica Acta* **2013**, *790*, 14. [CrossRef]

[14] Snabe, T.; Petersen, S. B. Application of infrared spectroscopy (attenuated total reflection) for monitoring enzymatic activity on substrate films. *Journal of Biotechnology* **2002**, *95*, 145. [CrossRef]

[15] Kumar, S.; Barth, A. Following Enzyme Activity with Infrared Spectroscopy. *Sensors* **2010**, *10*, 2626. [CrossRef]

[16] Barth, A. Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta (BBA) – Bioenergetics* **2007**, *1767*, 1073. [CrossRef]

[17] Ghose T. K.; Bisaria, V. S. Measurement of hemicellulase activities part 1: Xylanases. *Pure and Applied Chemistry* **1887**, *59*, 1739. [CrossRef]

[18] Vasconcelos, N. M.; Pinto, G. A. S.; Aragao, F. A. S. Determinação de açúcares redutores pelo ácido 3,5-dinitrosalicílico: histórico do desenvolvimento do método e estabelecimento de um protocolo para o laboratório de bioprocessos. *Boletim de pesquisa e desenvolvimento. Embrapa Agroindústria Tropical* **2013**, *88*, 29. [Link]

[19] Lloyd, J. B.; Whelan, W. J. An improved method for enzymic determination of glucose in the presence of maltose. *Analytical Biochemistry* **1969**, *30*, 467. [CrossRef]

[20] Savitzky, A.; Golay, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* **1964**, *36*, 1627. [CrossRef]

[21] Barnes, R. J.; Dhanoa, M. S.; Lister, S. J. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy* **1989**, *43*, 772. [CrossRef]

[22] Wold, S.; Sjostrom, M.; Eriksson, L. PLS-regression: A Basic Tool of Chemometrics. *Chemometrics and Intelligent Laboratory Systems* **2001**, *58*, 109. [CrossRef]

[23] Fearn, T. Assessing Calibrations: SEP, RPD, RER and R2. *NIR New* **2002**, *13*, 12. [CrossRef]

[24] König, J.; Grasser, R.; Pikor, H. Vogel, K. Determination of xylanase, β-glucanase, and cellulase activity. *Analytical and Bioanalytical Chemistri* **2002**, *374*, 80. [CrossRef]

[25] Decker S. R.; Adney W. S.; Jennings E.; Vinzant, T. B.; Himmel, M. E. Automated Filter Paper Assay for Determination of Cellulase Activity. *Applied Biochemistry and Biotechnology* **2003**, *107*, 689. [CrossRef]

[26] Lai, T. E.; Pullammanappallil, P. C.; Clarke, W. P. Quantification of cellulase activity using cellulose-azure. *Talanta* **2006**, *69*, 68. [CrossRef]

[27] Yu, X.; Liu, Y.; Cui, Y.; Cheng, Q.; Zhang, Z.; Lu, J. H.; Meng, Q.; Teng, L.; Ren, X. Measurement of filter paper activities of cellulase with microplate-based assay. *Saudi Journal of Biological Sciences* **2016**, *23*, S93. [CrossRef]

[28] Rinnan, A.; Berg, F.; Engelsen, S. B. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry* **2009**, *28*, 1201. [CrossRef]

[29] Xu, L.; Zhou, Y-P.; Tang, L-J.; Wu, H-L.; Jiang, J-H.; Shen, G-L.; Yu, R-Q. Ensemble preprocessing of near-infrared (NIR) spectra for multivariate calibration. *Analytica Chimica Acta* **2008**, *616*, 138. [CrossRef]

[30] Koonjah, S. S.; Beekharry, A.; Badaloo, M. G. H.; Henderson, C.; Dookun Saumtally, A. Evaluation of Near Infrared Spectroscopy for the Direct Analysis of Cane Quality Characters. *Universal Journal of Agricultural Research* **2019**, *7*, 169. [CrossRef]

[31] Souza, J. S.; Ferrão, M. F. Aplicações da espectroscopia no infravermelho no controle de qualidade de medicamentos contendo diclofenaco de potássio. Parte I: Dosagem por regressão multivariada. *Revista Brasileira de Ciências Farmacêuticas* **2006**, *42*, 437. [CrossRef]

[32] Abdi, H. Partial least squares regression and projection on latent structure regression (PLS Regression). *WIREs Computational Statatistics* **2010**, *2*, 97. [CrossRef]

[33] Viscarra Rossel, R. A.; Walvoort, D. J. J.; McBratney, A. B.; Janik, L. J.; Skjemstad, J. O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59. [CrossRef]

[34] Malley, D. F.; Martin, P. D.; Ben-Dor, E. Application in analysis of soils. *Near-infrared spectroscopy in agriculture*. Roberts, C. A.; Workman J.; Reeves, J. B. (Eds.). Agronomy, 44. ASA-CSSA-SSSA, Madison, WI, USA **2004**, *44*, 729. [CrossRef]