https://dx.doi.org/10.21577/0103-5053.20230020

J. Braz. Chem. Soc., Vol. 34, No. 8, 1078-1086, 2023

Article

Residual Lignocellulosic Biomass in the Production of HMF

Cristian A. Godoy,[a] *Patrícia Valderrama,*[b] *Andreia C. Furtado*[a] *and Marcela Boroski*[*,a]

[a]*Universidade Federal da Integração Latino-Americana (UNILA), Av. Tancredo Neves, 6731, 85867-970 Foz do Iguaçu-PR, Brazil*

[b]*Universidade Tecnológica Federal do Paraná (UTFPR), Via Rosalina Maria dos Santos, 1233, 87301-899 Campo Mourão-PR, Brazil*

5-Hydroxymethylfurfural (HMF) can be obtained from the dehydration of monosaccharides present in biomass, which helps add value to the raw material. The present work sought to study the relationship between the parameters of biomass characterization, including the content of reducing sugars, total soluble solids and the mid-infrared (MIR) and near-infrared (NIR) spectra, and the formation of HMF after synthesis. A thorough analysis was conducted in order to evaluate the residual agricultural biomass samples from sugarcane bagasse, corn stover, cassava branch, banana pseudostem and mango fruits after the drying (45 ºC maximum) and ball mill grinding pre-processing procedure. HMF synthesis was performed using 0.50 mol L^{-1} HCl and *n*-butanol as extracting solvent. The analytical quantification of HMF and furfural (co-product of synthesis) was performed using high-performance liquid chromatography coupled to diode-array detector (HPLC-DAD). The highest efficiency of the synthesis of HMF was observed for the mango fruits (55.0 kg ton^{-1}) and sugarcane bagasse biomass (47.5 kg ton^{-1}) samples. The application of data fusion and principal component analysis (PCA) allowed us to identify the samples with greater potential for HMF synthesis.

Keywords: biomass conversion, agricultural residues, validation, smoothing, FTIR-ATR, data fusion

Introduction

Brazil stands out as one of the leading producers of agricultural commodities on the world stage,[1] and the wastes generated in the production of these commodities are potential sources of renewable energy. The use of biomass as a source of raw material has enormous advantages; given the huge amount of biomass generated in the production of agricultural commodities, this biomass can be transformed into a wide range of commercial products, including renewable biofuel, products that help to neutralize carbon emission, low-cost biodegradable products, and so forth.[2] Furthermore, it is worth noting that the reutilization of agricultural residues from the agricultural industry does not lead to competition for sources of biomass used in food production.

The major residual industrial biomass produced in Brazil is sugarcane bagasse,[3] with values estimated of 690,832,317 tons of sugarcane produced in 2020.[4,5] Corn is the third most widely produced crop in Brazil, with the total area used for planting of 17,502,182 ha. The total planted area for cassava was 1,253,842 ha, making it the fourth most widely produced crop; and the total planted area for banana was 456,922 ha, making it the seventh most widely produced crop in the country.[5]

Sugarcane bagasse is used for the generation of energy either through combustion in boilers or via the utilization of the residue for the production of second-generation ethanol, where the process leads to the production of high volume of residual biomass.[3] With regard to the aforementioned crops, cassava branch, which represents the biggest volume of residue generated in the cultivation of cassava, is used for the feeding of animals on a small scale and for the planting of new areas to be cultivated; the cassava branch residue surplus is simply not put to any valuable use. Corn stover residue produced in corn plantations during the harvesting of corn grains is widely used in vegetation covering and for the feeding of animals.

*e-mail: marcela.boroski@unila.edu.br, marcelaboroski@yahoo.com.br
Editor handled this article: Andrea R. Chaves (Associate)

For each ton of banana produced, three tons of banana pseudostem are generated.[6] Based on the quantity of banana produced in 2019, more than 21 million tons of pseudostem were generated.[5] The residual biomass produced from the cultivation of banana is not utilized in the traditional way. Studies reported in the literature show that the residual biomass from the cultivation of banana is used for the production of second-generation ethanol[7] and biogas,[8] as well as for the fabrication of briquettes used for energy generation via combustion.[9]

Another residual biomass that is found in the western region of Paraná and which is derived from the afforestation practices implemented in the region is the biomass from mango fruits. To date, this biomass has not been put to any proper use; in fact, during the seasonal production period in the western region of Paraná, the streets can be found packed with mango fruits and a cleaning strategy is usually required to rid the streets of these fruits in order to avoid the breeding of insects.

Studies reported in the literature[2,6,10] have demonstrated the potential of the residual agricultural biomass investigated in this study for the generation of products that are commercially suitable for the chemical industry, and which have high added value. One of the ways in which residual agricultural biomass can be used for the generation of products of high added value is its utilization for the synthesis of 5-hydroxymethylfurfural (HMF).[10] HMF is a compound that occurs naturally in foods; this compound is derived from the dehydration of monosaccharides (glucose and fructose), in products subjected to high temperature heating treatments or to aging process. In terms of synthesis, HMF can be obtained from the elimination of three molecules from water during dehydration through acid catalysis; this process is more highly favored in the presence of fructose than in the presence of glucose.[11] Residues from some cultivars that occupy huge areas are particularly interesting as feedstock for HMF synthesis, such as sugarcane bagasse.[11-14] Also, from other lignocellulosic materials such as banana plant waste,[15,16] corn stover,[17,18] cassava waste,[19,20] and mango fruits.[21]

HMF is a compound of huge commercial interest in the chemical industry because it can be used as a precursor for the production of other chemical substances. HMF can be potentially converted into a wide range of chemical elements, such as 2,5-dimethylfuran (DMF) biofuel and 2,5-furandicarboxylic (FDCA) acid monomer, which enables the production of polyethylene furanoate (PEF) a polymer with mechanical properties similar to those of polyethylene terephthalate (PET). HMF can also be alternatively used for the production of tetrahydrofuran (THF), a solvent widely used in the chemical industry.[22]

Spectroscopic techniques are widely applied methods that enable to acquire relevant information regarding the chemical groups/classes present in biomass.[23] The application of spectroscopic techniques in the infrared region, in the mid-infrared (MIR) or near-infrared (NIR) region, is useful because it offers a clear elucidation of the chemical structure, providing relevant characteristics regarding the chemical groups present in the biomass.[24] Spectral modifications related to the content of the monomers of each carbohydrate can be monitored through the application of the attenuated total reflectance (ATR) technique in the MIR region, and this can be used as an alternative characterization method for the analysis of biomass sources which are found to be highly promising for the synthesis of HMF. These spectroscopic techniques have a wide range of advantages which include the following: (*i*) low cost of analysis; (*ii*) ability to perform direct analysis of raw samples; (*iii*) requires the use of minimal sample preparation via milling and standardization of particle size. Chemometric tools are commonly used in spectral analysis to highlight the characteristics of each sample such as the presence of certain monomers or other structures related to the sample composition. In this sense, the application of data fusion allows one to obtain more information based on an exploratory investigation through principal component analysis (PCA);[25] the implementation of this analytical methodology provides clear information in terms of the production of HMF by a certain type of biomass.

The present work sought to investigate the relationship between the characteristic parameters of biomass, such as infrared spectra, monitored by Fourier-transform infrared spectroscopy coupled to attenuated total reflectance (FTIR-ATR) and Fourier-transform near-infrared spectroscopy (FT-NIR), the content of reducing sugars and soluble solids, and the production of HMF after acid catalysis. The results obtained from this analysis enabled us to effectively characterize the types of biomass investigated in the present study in order to have a thorough understanding regarding the potential of the biomass in terms of the synthesis of HMF.

Experimental

Sampling

The sources of residual agricultural biomass used in this study were as follows: sugarcane bagasse (SB), banana pseudostem (BP), corn stover (CS), cassava branch (CB), and mango fruits (MF) (Figure S1, Supplementary Information (SI) section).

The biomass samples were obtained from the São Miguel do Iguaçu municipality, which is located in the

western region of Paraná, Brazil. The biomass samples were obtained taking into account the seasonal cultivation periods. The sugarcane bagasse biomass samples were acquired from a sugarcane plantation situated at the Global Positioning System (GPS) coordinates 25.287934, −54.269090, in August 2020. The corn stover biomass samples were acquired from a corn plantation situated at the GPS coordinates −25.249578, −54.361730, in July 2020. The banana pseudostem and cassava branch biomass samples were obtained from plantations located close to the corn plantation in August and September 2020, respectively. Finally, the entire mango fruit biomass samples were obtained from the urban area of São Miguel do Iguaçu municipality in August 2020 (based on the fruit maturation).

Sample preparation

The drying of the samples was performed in two steps to ensure the preservation of the biomass. The first step involved drying the samples naturally for one week at room temperature, storing them in the shade and in an open space. The samples were then ground in a disintegrator (MA758, Marconi, Piracicaba, Brazil). Subsequently, the samples *in natura* were subjected to drying at 45 ºC for 72 h in a sterilization and drying oven (Luca-80/480, Lucadema, São José do Rio Preto, Brazil).

The milling of the biomass was carried out based on the methodology described by Flores-Velázquez *et al*.[15] Several adjustments were made in the grinding configuration until obtaining a satisfactory milling procedure which yielded a particle size less than 180 µm. Under this adapted configuration, 16 steel balls of 10 mm and 80 steel balls of 5 mm were used for 8.75 g of biomass placed in a steel vat with a volume of 500 mL (Figure S2, SI section). The steel vat containing the steel balls and the biomass was fixed onto the vat support of a ball mill (PM 100, Retsch, Haan, Germany). The material was subjected to milling cycles for 30 min at a rotation of 400 rpm. Once the milling of each biomass was completed, the biomasses were sieved in a stainless steel particle size analysis sieve that had a 180 µm opening.

Characterization

The sample moisture was determined in an oven (LT 100 CR, LimaTec, Guarulhos, Brazil) at 105 ºC for 6 h until a constant weight was obtained.[26,27] The quantification of reducing sugars was performed using classical methodologies.[26,28] The analysis of total soluble solids was performed using a benchtop refractometer (2WAJ, Biobrix, São Paulo, Brazil) with reading in °Brix directly on the equipment.[29] Clear samples obtained after the hydrolysis of the biomass were used, as described in the procedure for the determination of reducing sugars, including the clarification with Carrez I (potassium ferricyanide, ≥ 99%, Dinâmica, Indaiatuba, Brazil) and Carrez II (zinc sulfate, ≥ 99%, Reatec, Colombo, Brazil) solutions, filtration and centrifugation (Rotina 380, Hettich, Tuttlingen, Germany).

Synthesis and analysis of 5-hidroxymethylfurfural (HMF)

For the synthesis of HMF, a methodology applied in previous studies[11,30] was used for the choice of the reaction medium; under this methodology, the optimal condition of synthesis was obtained from the application of 0.50 mol L^{-1} hydrochloric acid (HCl 37% (m/m) PA; Neon, Suzano, Brazil) as homogeneous catalyst. The synthesis of the biomass was carried out in triplicate. Heating blankets were used for the heating of the flasks, and an ultrathermostat bath was used for the cooling of the condensation columns (Solab/sl 152). The mixture was heated under reflux at 85 ºC for 60 min. At the end of this period, 3.0 g of NaCl (99.5%, Exodo, Sumaré, Brazil) and 20 mL of *n*-butanol (99.4%, Synth, Diadema, São Paulo, Brazil) were added into the flask, and the mixture was subjected to heating at 155 ºC for 90 min. A 600 mm Allihn condenser was connected to each flat-bottomed flask, and the condenser was connected to a bubbling system. After the cooling down of the reaction medium at room temperature, the mixture was subjected to filtration using a 250 mL kitassate under vacuum and the aqueous phase was then separated from the organic phase using a 250 mL separating funnel. The synthesis of the biomass was carried out in triplicate.

The separation and quantification of the analytes of interest were carried out using high-performance liquid chromatography (HPLC) system, (DionexUltimate 3000, Thermo Fisher Scientific, Germering, Germany), coupled to a diode-array detector (DAD 3000) detailed in our previous studies.[11,30]

The separation was made in an ACE 5 C18 column (batch no. V13-7473) (250 mm × 4.6 mm; particle size of 5 µm; particle porosity of 110 °A), kept at 30 °C. The injection volume was 20 µL. The diode array detector recorded the spectra in the range from 200 to 400 nm, with detection of the analytes at specific wavelengths of 263, 277, and 285 nm for FDCA, furfural (FF), and HMF, respectively. The elution of the analytes (FDCA, HMF, and FF) were performed using a 0.01 mol L^{-1} trisodium citrate solution (at pH 2.5) as the mobile phase. The analytical curves were constructed with HMF (purity ≥ 99.0%, Sigma-Aldrich, Saint Louis, Missouri, USA), FF (purity of 99.0%, Sigma-Aldrich, Saint Louis, Missouri,

USA), and 2,5-furandicarboxylic acid (FDCA, purity 97.0%, Sigma-Aldrich, Saint Louis, Missouri, USA) as the internal standard.

Synthesis sample at 10 μL was added to 890 μL of ultrapure water and 100 μL of FDCA at 2.00×10^{-1} mmol L^{-1}, followed by homogenization. The samples were filtered through 0.22 μm polytetrafluoroethylene (PTFE) hydrophilic syringe filters and were then kept at –20 °C, prior to the chromatographic analyses. For each of the three synthesis replicates, three sample replicates in the aqueous phase and three sample replicates in the organic phase were prepared for injection into the HPLC system.

Spectroscopy in the infrared region

The biomass samples with particle size of 180 μm were analyzed in the near-infrared (NIR) region using Shimadzu equipment with Fourier-transform infrared spectroscopy coupled to an attenuated total reflectance accessory (IRAffinity-1, Shimadzu, Kyoto, Japan) in the range of 650 to 3500 cm^{-1}, with 8 cm^{-1} resolution. The NIR spectra of these biomass samples and the biomass samples obtained from the acid hydrolysis related to the preparation of reducing sugars and the organic phases of synthesis were recorded in the microNIR (JDSU, Milpitas, USA) equipment in the range of 950 to 1650 nm, under the resolution of 8 cm^{-1}.

Chemometric analysis

The NIR spectra for the biomass, organic phase and acid hydrolysis samples were pre-processed separately using the Savitzky-Golay (savgol) smoothing algorithm (first derivative, first-order polynomial, 7-point window).[31] After the pre-processing of the data, the spectra were concatenated, along with the results obtained from the analysis of moisture, reducing sugars, Brix, and the concentrations of HMF and FF, in low-level data fusion with the application of PCA. All the calculations were performed using the Matlab software R2007b[32] along with the analytical tools from PLS-Toolbox 5.1.[33] The objective of this analysis was to explore the ability of NIR spectroscopy to highlight the potential of the biomass investigated in this study for the synthesis of HMF.

Results and Discussion

Analysis and synthesis of biomass

A thorough analysis was performed in order to evaluate the characteristics of the ground biomass investigated in this study. Among the characteristics investigated included the following: moisture content, content of reducing sugars, and total soluble solids (ºBrix). Table 1 presents the results obtained from this analysis along with the quantification data obtained for HMF and FF in the aqueous and organic phases of the synthesis.

The moisture of the samples employed in this study varied between 6.81% (sugarcane bagasse) and 9.95% (corn stover). This low moisture content observed for all the biomass samples provides greater chemical and physical stability during storage. In a recent study published in the literature, Kumar *et al.*[34] employed different types of biomass as raw material for the synthesis of HMF; the types of biomass they employed had moisture content less than 10%. Borges *et al.*[35] also employed biomass abundant in the region where they conducted their study for the synthesis of HMF; the biomass they employed had moisture content less than 7.11%. For the biomass samples investigated in our present study, the contents of reducing sugars obtained in the samples ranged between 17.96% (for biomass from mango fruits), and 43.77% (for corn stover). The quantification analysis of reducing sugars provides us with useful information regarding the availability of monosaccharides in the biomass samples. The results obtained from the quantification of total soluble solids, measured in ºBrix, did not show any significant differences between the biomass samples investigated (Table 1).

Table 1. Characterization of the biomass samples and contents of HMF and FF obtained after the syntheses

Sample	Moisture / %	Reducing sugars / %	ºBrix / %	[HMF] AP / (mmol L^{-1})	[HMF] OP / (mmol L^{-1})	[FF] AP / (mmol L^{-1})	[FF] OP / (mmol L^{-1})
SB	6.81 ± 0.02	35.78 ± 2.98	8.63 ± 0.18	3.67 ± 0.96	16.10 ± 2.40	0.59 ± 0.14	5.64 ± 2.05
CS	9.95 ± 0.29	43.77 ± 4.76	9.00 ± 0.00	0.53 ± 0.15	1.66 ± 0.50	0.59 ± 0.25	4.37 ± 1.75
CB	9.56 ± 0.07	30.37 ± 1.58	9.00 ± 0.00	1.20 ± 0.01	3.25 ± 0.16	0.15 ± 0.04	1.12 ± 0.14
MF	8.27 ± 0.03	17.96 ± 1.37	9.00 ± 0.00	4.94 ± 0.62	18.10 ± 1.79	0.08 ± 0.02	0.78 ± 0.17
BP	6.91 ± 0.10	20.75 ± 0.86	9.00 ± 0.00	0.89 ± 0.10	2.26 ± 0.61	0.09 ± 0.04	0.87 ± 0.41

Results obtained in triplicate and presented as mean ± standard deviation. HMF: 5-hydroxymethylfurfural; FF: furfural; AP: aqueous phase; OP: organic phase; SB: sugarcane bagasse; CS: corn stover; CB: cassava branch; MF: mango fruits; BP: banana pseudostem.

The HMF concentration values obtained in the syntheses ranged between 0.53 mmol L^{-1} for the aqueous phase of the corn stover sample and 18.10 mmol L^{-1} for the organic phase of the mango fruits sample. Regarding the FF concentration generated as a co-product of synthesis, the concentration values obtained were in the range of 0.08 mmol L^{-1} for the synthesis involving the use of biomass from mango fruits in aqueous phase to 5.64 mmol L^{-1} for the synthesis involving the use of sugarcane bagasse biomass in the organic phase.

The biomass from corn stover exhibited the lowest efficiency in terms of the production of HMF; the corn stover biomass generated the least amount of HMF but presented the highest fraction of FF produced among all the different types of biomass investigated. This result makes the biomass from corn stover unsuitable for use in the production of HMF precursor due to the elevated energy costs involved in the purification of the product of synthesis.

The biomass from mango fruits produced high amount of HMF and much lower amount of FF (Table 2), making it the biomass with the lowest energy costs when it comes to the purification of the product of synthesis.

Table 2. Production of HMF *per* ton of biomass

Sample	HMF *per* biomass / (kg ton^{-1})	FF *per* biomass / (kg ton^{-1})
SB	47.5	11.7
CS	5.2	9.2
CB	10.5	2.4
MF	55.0	1.6
BP	7.4	1.8

Mass of HMF (5-hydroxymethylfurfural) and (FF) furfural, taking into account the sum of the compound masses in aqueous and organic phases. Reference: ton of biomass on a dry basis. SB: sugarcane bagasse; CS: corn stover; CB: cassava branch; MF: mango fruits; BP: banana pseudostem.

Characterization of biomass by FTIR-ATR

The biomass samples were characterized by FTIR-ATR spectroscopy, and the spectra obtained are shown in Figure 1. It is possible to observe important peaks in distinguishing among different biomass around 1300, 1600, 1700, and 3000 cm^{-1}.

Biomass from mango fruits (MF) exhibited the highest efficiency in terms of the synthesis of HMF; this biomass sample presented transmittance intensity bands close to the region around the wavelengths of 1700 and 3000 cm^{-1}. The biomass samples from corn stover (CS) and banana pseudostem (BP) presented the lowest efficiency in the synthesis of HMF; these samples exhibited bands of higher transmittance intensity in regions close to the wavelengths of 1600 and 1300 cm^{-1}.

To interpret the transmittance spectra of the biomass samples investigated in this study, the characterization graph of each functional group was evaluated. The following conclusions were drawn based on the results obtained from this analysis: (*i*) the bands obtained for the MF biomass are related to strong absorption of alkenes and medium or weak absorption of carbonyl in the region around the wavelength of 1700 cm^{-1}, as well as to medium and weak absorption of alkenes and aromatic groups, respectively, in the region around the wavelength of 3000 cm^{-1}; (*ii*) with regard to the CS and BP biomass samples, the bands in the region around the wavelength of 1300 cm^{-1} correspond to strong absorption of sulfur dioxide, strong or medium absorption of the nitro functional group, and weak absorption of alkenes, while the bands in the region around the wavelength of 1600 cm^{-1} are associated with strong absorption of carbonyl and nitro functional group and weak absorption of aromatic groups.

Another interesting point to note regarding the spectra of the biomass samples is the band formed after the wavelength of 1000 cm^{-1}; this region is characterized by the absorptions of hydroxyl, oxide and sulfur dioxide. The fact that the samples analyzed are biomasses largely composed of cellulose, one expects to see high absorption of hydroxyl in the region, due to the structure of cellulose, which is a glucose polymer (glucose molecule has six hydroxyls in its structure).

The bands in the wavelength region of 1300 and 1600 cm^{-1}, which are found in greater intensity in the biomass samples with lower efficiency in the synthesis of HMF, represent the chains that contain alkenes and sulfur dioxide and nitro radicals. These three groups do not take part in the primary reaction process, and can reduce the reaction speed and efficiency, acting as inhibitors; alternatively, due to the occurrence of secondary reactions, considering that the three groups are highly reactive, they can act as compounds that are potentially competing for reaction.

Data fusion and PCA

Figure 2 shows the NIR spectra obtained after the pre-processing procedure. As can be observed, in the NIR region, the water presents a first overtone O–H stretch band at 1450 nm. However, other compounds with O–H bonds, including carbohydrates and lignin, can also be associated with the region.[36]

Furthermore, the NIR spectra of the biomass samples present a vibrational band which corresponds to the alkyl

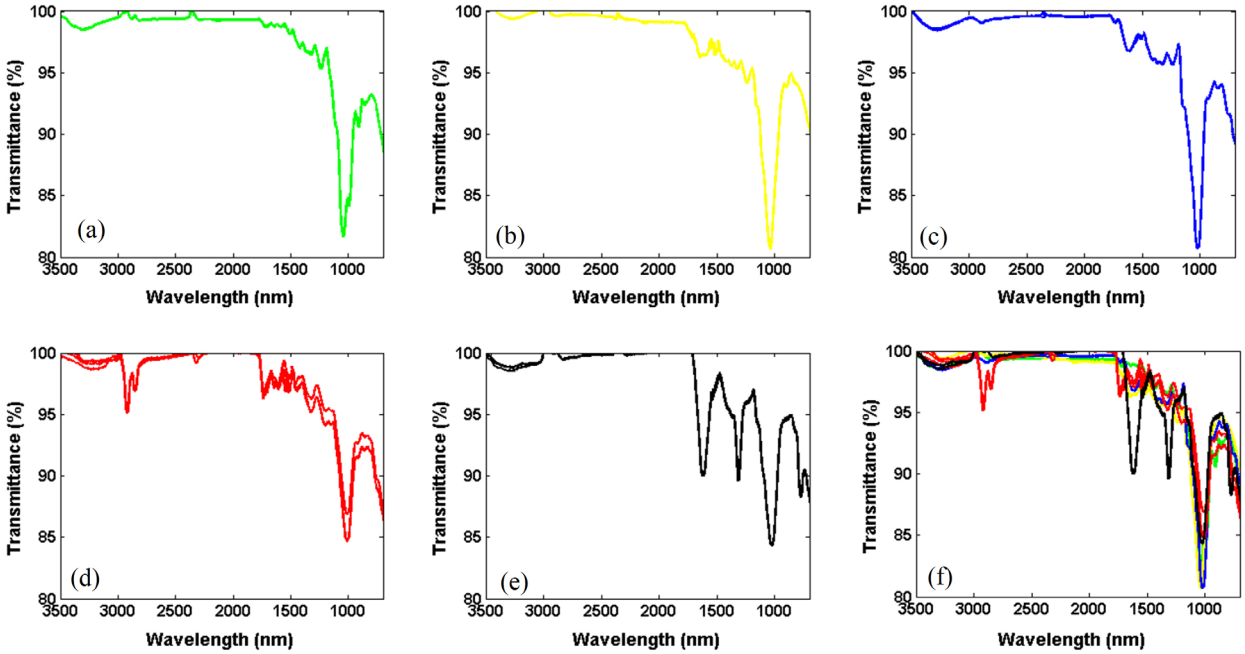

Figure 1. MIR spectra obtained for the following biomass samples: (a) SB (–); (b) CS (–); (c) CB (–); (d) MF (–); (e) BP (–); (f) all biomass samples together.

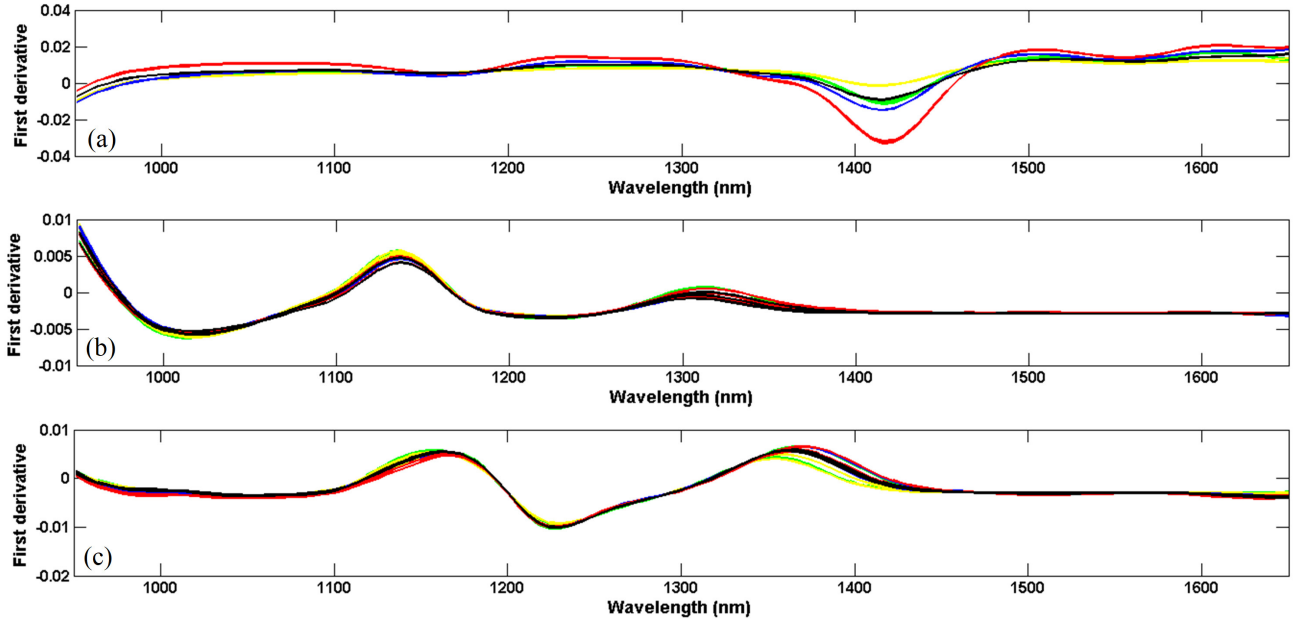

Figure 2. NIR spectra after the pre-processing procedure. (a) Biomass; (b) aqueous phase; (c) organic phase; (–) SB; (–) CS; (–) CB; (–) MF; (–) BP.

groups (C–H*n*) at 1190 nm (second overtone of the C–H stretch). Lignin exhibits specific bands that correspond to C–H or O–H aromatic groups at 950 and 1100 nm.[36,37] Nonetheless, the characteristic spectral overlap of the NIR region points to the difficulty of carrying out a direct analysis of the spectra without the aid of multivariate tools. On the other hand, the conventional analytical techniques require considerable amount of time and are extremely costly to perform; apart from that, these techniques generate residues that need to be disposed of in a proper fashion. In view of that, a thorough analysis of the potential of different biomass samples in the synthesis of HMF can be conducted in a more rapid way and at a relatively cheaper cost using NIR spectroscopy.

The samples from acid hydrolysis and the organic phase of the synthesis were evaluated by NIR spectroscopy. These results were concatenated along with those shown in Table 1, creating low-level data fusion (fusion of pre-processed individual matrices), followed by the application of the principal component analysis (PCA).

The NIR spectra (after the first derivative), along with the results shown in Table 1 (after auto-scaling), were normalized through maximum intensity prior to concatenation.

The PCA, applied for the analysis of the merged data, centered on the mean and two principal components (PCs), was able to capture 78.67% of the explained variance. Based on the scores in Figure 3, one can see a clear distinction between the raw materials.

The results obtained from the joint analysis of scores and loadings (Figure 4) showed that SB and CS are distinguished by the positive scores of PC1 due to the content of moisture, reducing sugars and the concentration of furfural in the aqueous and organic phases. The PC1 positive loadings related to the NIR spectra show the following: (*i*) the important region in the biomass samples is between the wavelengths of 1300 to 1500 nm; (*ii*) the important regions for the acid hydrolysis samples are between the wavelengths of 950 to 1300 nm and between 1320 to 1650 nm; and (*iii*) the important regions for the organic phase samples are around the wavelengths of 1140, 1210 and above 1600 nm. All these regions are the spectral regions that are mostly associated with the contents of moisture, reducing sugars and furfural. MF and BP in the negative part of PC1 were found to present similar results for the parameters mentioned above.

The results obtained from the analysis of PC2 showed that MF is situated in the most positive part of this PC. In fact, this biomass sample presented the highest concentration of HMF in both aqueous and organic phases, and these concentrations were found to diminish for the

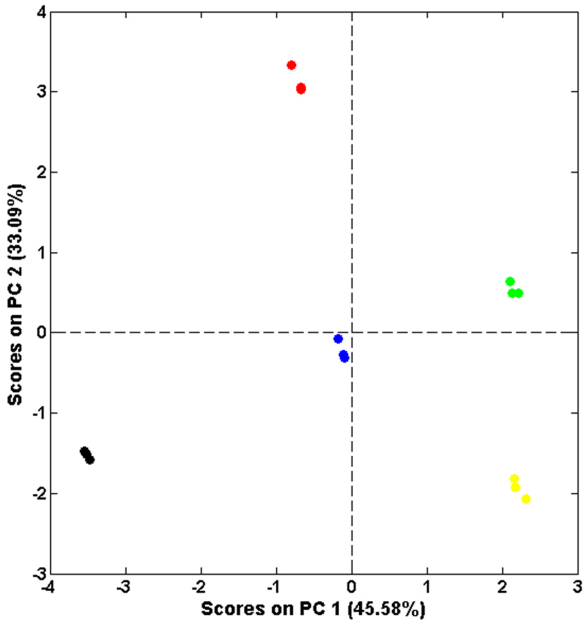

Figure 3. PCA scores on PC2. (•) SB; (•) CS; (•) CB; (•) MF; (•) BP.

biomass samples based on the decrease in the PC2 scores: MF > SB > CB > BP > CS. The PC2 loadings related to the NIR spectra showed that the spectral regions related positively to higher concentrations of HMF were as follows: (*i*) the wavelength region between 950-1100 nm for the biomass samples; (*ii*) the region around 1320 nm for the acid hydrolysis sample; and (*iii*) the region around 1390 nm for the organic phase sample.

The application of data fusion and PCA enabled us to identify the spectral regions in the samples of biomass,

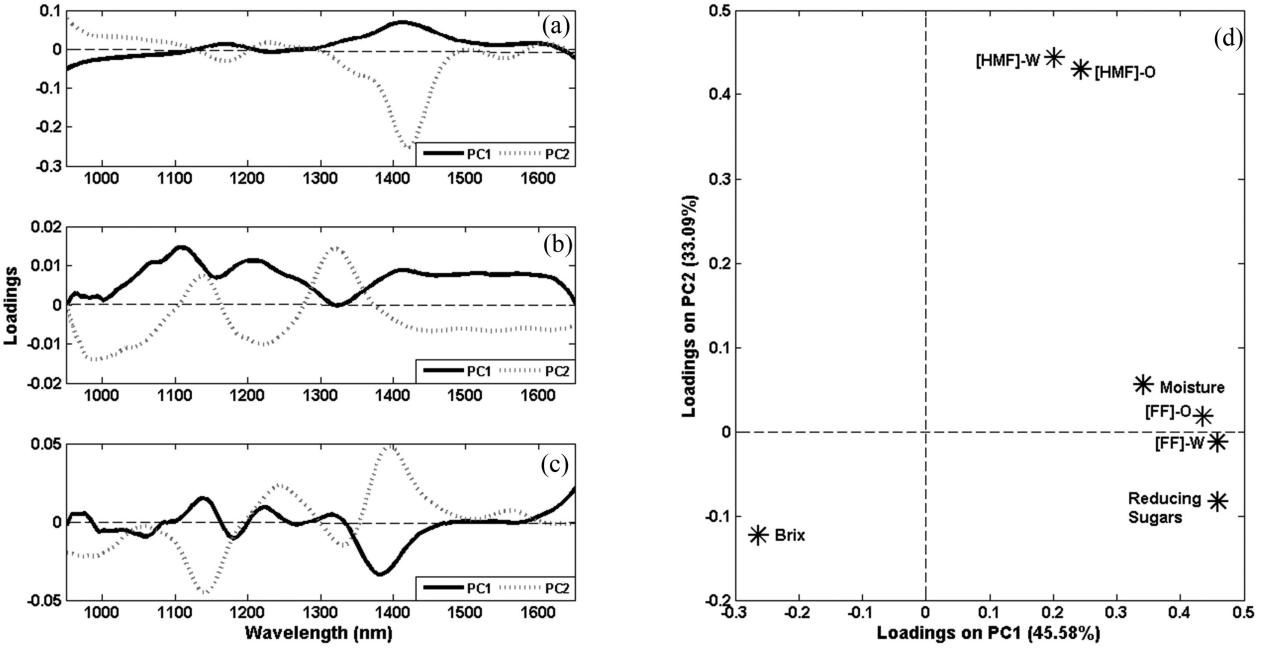

Figure 4. PCA loadings. (a) NIR for biomass samples; (b) NIR for acid hydrolysis samples; (c) NIR for organic phase samples; (d) conventional analyses.

acid hydrolysis and organic phase that indicated possible potential for the synthesis of HMF.

Conclusions

The results obtained from the analysis of the potential of different samples of biomass for the synthesis of HMF showed that the biomass sample from mango fruits (MF) exhibited the highest efficiency in HMF production, followed by the biomass samples from sugarcane bagasse (SB), cassava branch (CB), banana pseudostem (BP) and corn stover (CS), in this order. In terms of the scale of HMF production per ton of biomass, the production rate ranged between 5 and 55 kg ton^{-1}; this shows that the synthesis route employed provides one with the possibility of valorizing the residual biomass produced by the agricultural industry.

The results obtained from the FTIR-ATR analyses conducted in the mid-infrared (MIR) region showed the transmittance bands related to the biomass samples; among these bands included the bands in the regions of inhibiting domain or those that competed for reaction, and bands in the regions that were associated with functional groups which are typically related to the reagents used in the synthesis. With regard to the analysis in the near-infrared region, the application of multivariate tools allowed us to concatenate the NIR spectra, along with the results obtained from the characterization of the ground biomass samples and the data related to the quantification of HMF and FF in the aqueous and organic phases of the synthesis; this paved the way toward the conduct of principal component analysis. Based on the results of these analyses, one was able to observe a clear distinction between the raw materials investigated in this study and match the spectral regions of the biomass samples with their potential for the synthesis of HMF.

Supplementary Information

Supplementary data (Figures S1 and S2) are available free of charge at http://jbcs.sbq.org.br as PDF file.

Acknowledgments

Marcela Boroski is grateful to the National Council for Scientific and Technological Development (Conselho Nacional de Tecnologia e Desenvolvimento) (CNPq) - Process code number: 420065/2016-4, and the "Triple Agenda" Institutional Program (UNILA) for the financial assistance provided in support of this research. Patrícia Valderrama would also like to express her gratitude to CNPq (process No.: 306606/2020-8) for the financial assistance granted in support of this work. The authors are extremely grateful to the Center for Renewable Energies (Centro de Energias Renováveis-PTI) for the assistance with research equipment and facilities and to the LEIMAA Research Group for the technical support.

References

1. Food and Agriculture Organization of the United Nations (FAOSTAT); *The Future of Food and Agriculture - Trends and Challenges*; FAO: Rome, 2017. [Link] accessed in January 2023
2. Aristizábal, M. V.; Gómez, P. Á.; Cardona, C. A. A.; *Bioresour. Technol.* **2015**, *196*, 480. [Crossref]
3. Carpio, L. G. T.; de Souza, F. S.; *Renewable Energy* **2017**, *111*, 771. [Crossref]
4. Instituto Brasileiro de Geografia e Estatística (IBGE); *Indicadores IBGE Estatística da Produção Agrícola*; 2016. [Link] accessed in January 2023
5. Instituto Brasileiro de Geografia e Estatística (IBGE); *Indicadores IBGE Levantamento Sistemático da Produção Agrícola Estatística da Produção Agrícola*; 2020. [Link] accessed in January 2023
6. Souza, O.; Federizzi, M.; Coelho, B.; Wagner, T. M.; Wisbeck, E.; *Rev. Bras. Eng. Agricola Ambient.* **2010**, *14*, 438. [Crossref]
7. Gonçalvez Filho, L. C.: *Utilização do Pseudocaule de Bananeira como Substrato da Fermentação Alcoólica: Avaliação de Diferentes Processos de Despolimerização*; MSc Dissertation, Universidade da Região de Joinville (Univille), Joinville, Brazil, 2011. [Link] accessed in January 2023
8. Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA); *Produção de Biogás a Partir do Líquido do Pseudocaule da Bananeira*; 2012. [Link] accessed in January 2023
9. Athayde, C. S.: *Análise dos Resíduos Gerados pela Bananicultura como Possível Fonte de Geração de Energia*; MSc Dissertation, Federal University of Minas Gerais, Belo Horizonte, 2014. [Link] accessed in January 2023
10. Dulie, N. W.; Woldeyes, B.; Demsash, H. D.; Jabasingh, A. S.; *Waste Biomass Valorization* **2021**, *12*, 531. [Crossref]
11. Godoy, C. A.; Simião, L. M.; Toci, A. T.; Cordeiro, G. A.; do Amaral, B.; Peralta-Zamora, P.; Valderrama, P.; Boroski, M.; *BioEnergy Res.* **2020**, *13*, 737. [Crossref]
12. Li, M.; Jiang, H.; Zhang, L.; Yu, X.; Liu, H.; Yagoub, A. E. A.; Zhou, C.; *Ind. Crops Prod.* **2020**, *149*, 112361. [Crossref]
13. Gomes, G. R.; Rampon, D. S.; Ramos, L. P.; *J. Braz. Chem. Soc.* **2018**, *29*, 1115. [Crossref]
14. Ghorbannezhad, P.; Shen, G.; Ali, I.; *Biomass Convers. Biorefin.* **2022**. [Crossref]
15. Flores-Velázquez, V.; Córdova-Pérez, G. E.; Silahua-Pavón, A. A.; Torres-Torres, J. G.; Sierra, U.; Fernández, S.; Godavarthi, S.; Ortiz-Chi, F.; Espinosa-González, C. G.; *Fuel* **2020**, *265*, 116857. [Crossref]

16. de Souza, P. K.; Sellin, N.; Souza, O.; Marangoni, C.; *Chem. Eng. Trans.* **2013**, *32*, 1141. [Crossref]
17. Yang, T.; Chen, D.; Li, W.; Zhang, H.; *Mol. Catal.* **2021**, *515*, 111920. [Crossref]
18. Qing, Q.; Wu, W.; Tao, X.; Ma, Z.; He, W.; Tao, Y.; Wang, L.; *Ind. Crops Prod.* **2022**, *188*, 115681. [Crossref]
19. Manurung, R.; Winda, O.; Silalahi, H.; Siregar, A. G. A.; *Int. J. Eng. Res. Afr.* **2021**, *53*, 190. [Crossref]
20. Daengprasert, W.; Boonnoun, P.; Laosiripojana, N.; Goto, M.; Shotipruk, A.; *Ind. Eng. Chem. Res.* **2011**, *50*, 7903. [Crossref]
21. Muñiz-Valencia, R.; Portillo-Pérez, G.; Ceballos-Magaña, S. G.; Cortés-Quintero, G. C.; Nava-García, A. Y.; Dumont, M.-J.; Pineda-Urbina, K.; *Biomass Convers. Biorefin.* **2022**, *12*, 5145. [Crossref]
22. Galaverna, R.; Pastre, J. C.; *Rev. Virtual Quim.* **2017**, *9*, 248. [Crossref]
23. Lee, J.-W.; Ha, M.-G.; Yi, Y.-B.; Chung, C.-H.; *Carbohydr. Res.* **2011**, *346*, 177. [Crossref]
24. Song, J.; Jia, Y.-X.; Su, Y.; Zhang, X.-Y.; Tu, L.-N.; Nie, Z.-Q.; Zheng, Y.; Wang, M.; *Polymers* **2020**, *12*, 348. [Crossref]
25. Gonçalves, T. R.; Rosa, L. N.; Torquato, A. S.; da Silva, L. F. O.; Março, P. H.; Gomes, S. T. M.; Matsushita, M.; Valderrama, P.; *Food Anal. Method.* **2019**, *13*, 86. [Crossref]
26. *Normas Analíticas do Instituto Adolfo Lutz*, 2nd ed.; Instituto Adolfo Lutz: São Paulo, 1976. [Link] accessed in January 2023
27. Association of Official Analytical Chemists (AOAC); *Official Methods of Analysis of AOAC International*, 14th ed.; AOAC: Washington, 1984. [Link] accessed in January 2023
28. DuBois, M.; Gilles, K. A.; Hamilton, J. K.; Rebers, P. A.; Smith, F.; *Anal. Chem.* **1956**, *28*, 350. [Crossref]
29. Chaicouski, A.; da Silva, J. E.; da Trindade, J. L. F.; Canteri, M. H. G.; *Braz. J. Agrib. Prod.* **2014**, *16*, 43. [Crossref]
30. Godoy, C. A.; Valderrama, P.; Furtado, A. C.; Boroski, M.; *MethodsX* **2022**, *9*, 101774. [Crossref]
31. Savitzky, A.; Golay, M. J. E.; *Anal. Chem.* **1964**, *36*, 1627. [Crossref]
32. *MATLAB*, version R2007b; The MathWorks, Inc.: Natick, MA, USA, 2007.
33. *PLS Toolbox*, version 5.2; Eigenvectors Research: Manson, WA, USA, 2009.
34. Kumar, A.; Chauhan, A. S.; Shaifali; Das, P.; *Cellulose* **2021**, *28*, 3967. [Crossref]
35. Borges, M. S.; Barbosa, R. S.; Rambo, M. K. D.; Rambo, M. C. D.; Scapin, E.; *Biomass Convers. Biorefin.* **2020**, *12*, 3055. [Crossref]
36. Rambo, M. K. D.; Amorim, E. P.; Ferreira, M. M. C.; *Anal. Chim. Acta* **2013**, *775*, 41. [Crossref]
37. Shenk, J. S.; Workman, J. J.; Westerhaus, M. O. In *Handbook of Near-Infrared Analysis*, vol. 35, 3rd ed.; Burns, D. A.; Ciurczak, E. W., eds.; CRC Press: Florida, USA, 2008.

Submitted: October 18, 2022
Published online: February 13, 2023