Estimation of Xylose Recovery from Lignocellulosic Biomass

M T A P Kresnowati¹,²*, D C Januardi¹ and S V Utomo¹

¹Department of Chemical Engineering, Faculty of Industrial Technology, Bandung Institute of Technology, Indonesia
²Food and Biomass Processing Technology Research Group, Faculty of Industrial Technology, Bandung Institute of Technology, Indonesia

*Corresponding Author’s E-mail: kresnowati@che.itb.ac.id

Abstract. Lignocellulosic materials are potential raw materials for (bio)chemical industries due to their abundance. Its hemicellulosic content, for example, can be hydrolysed to xylose and later converted to various valuable biochemical products, e.g. xylitol. Due to the variability in characteristics and composition of the lignocellulosic materials, however, thorough research is required before the utilization of each type of lignocellulosic materials. This paper presents the development of an empirical model to estimate the yield of xylose from various lignocellulosic materials. A comprehensive literature study was conducted to build lignocellulosic database, in which the yields of xylose from various lignocellulosic materials that were processed by using different pretreatment condition were mapped. An empirical model was developed to establish a correlation between the type of lignocellulosic materials as well as the pretreatment operation condition (severity factor) and the yield of xylose. Several correction factors, such as biomass composition, lignin structure, and the succeeding hydrolysis process have been proposed to improve its accuracy.

1. Introduction
As an agricultural country, Indonesian produce massive amount of agricultural and plantation of biomass waste. In 2010, Indonesia produces 145 MT of biomass waste every day [20]. Biomass is classified as a renewable energy source with low economic value. In addition, following the biorefinery concept, biomass can also be used as the raw materials to produce high value-added products. However, the utilization of biomass has not been maximized.

Lignocellulosic biomass is generally composed of starch, cellulose, hemicellulose, and lignin [1]. The potentials of cellulose or lignin components in biomass have been thoroughly studied. However, the potential of hemicellulose utilization has been overlooked. Xylan, the biopolymers backbone of hemicellulose, is known as the second most abundant biopolymer in nature [23]. Xylan can be processed to recover the monosaccharide xylose that can be further converted into xylitol.

The processing of lignocellulosic materials to recover monosaccharides includes pretreatment and hydrolysis. The pretreatment step is aimed to release the polysaccharides which bind closely to the lignin contained in the biomass. This first step can be conducted either chemically with acid or physically through hydrothermal whereas the hydrolysis which can be done using enzymes or acids [27]. To recover xylose from the lignocellulosic biomass, first the bonds between hemicellulose and lignin need to be broken and further the free hemicellulose needs to be extracted and hydrolysed. For xylitol production, the produced hydrolysate can be directly fermented by a suitable microbial platform to produce xylitol such as Candida utilis or Debaryomyces hansenii [9].
One of the critical factors in producing xylitol from biomass is the percentage of xylose that can be recovered from biomass. The later is affected by the operating conditions applied in the pretreatment and biomass hydrolysis [15,21]. However, the suitable pretreatment operating conditions may also vary with the type of lignocellulosic biomass used as the raw materials [28]. Thereby before using a particular biomass a thorough research needs to be conducted to obtain the optimum operation condition for the biomass processing.

The purpose of this study was to collect a solid database on lignocellulosic biomass processing to obtain xylose and to map those data to examine the effects of pretreatment type as well as the related operating conditions on xylose recovery. In particular the widely available data on biomass waste from corn processing, oil palm processing and sugar cane processing were used to build a model. A model would be development to enable further estimation of xylose recovery from any type of lignocellulosic biomass materials at a predesigned pretreatment condition. Thereby reducing the wet laboratory experiments that need to be conducted in an attempt for biomass utilization.

2. Methodology
This research used secondary data which was obtained from an extensive literature study on research papers on the processing and utilization of lignocellulosic biomass: corn fiber [22], corn cob [4,6,14,25], palm oil fruit bunch [5,18,19], corn stover [5], and sugarcane bagasse [2,26]. The research was focused on biomass waste from corn processing: corn cob, corn stover, and corn fiber; biomass waste from oil palm processing: oil palm empty fruit bunches (EFB); and biomass waste from sugar cane processing: bagasse; which were selected due to their high hemicellulose content and their high availability in Indonesia. Later, data on wheat straw was also used for model validation.

The data was processed to obtain the severity factor (SF) and the combined severity factor (CSF) of the applied lignocellulosic biomass pretreatment operating conditions, following eq. (1) and eq. (2). In particular, SF was applied for the datasets obtained from the hydrothermal pretreatment process whereas CSF was applied for the datasets obtained from other types of pretreatment process, for example weak acid hydrolysis.

\[ SF = t \times e^{\frac{T-100}{1475}} \]  

In which \( t \) is the duration of the pretreatment process [minute], and \( T \) is the temperature of the pretreatment (°C) [21]. Meanwhile,

\[ \log CSF = \log SF - \text{pH} \]

In which the acidity of the pretreatment solution is expressed as the \( \text{pH} \) [Ruiz].

Model development, that was model parameters estimation and model simulation, was conducted in MATLAB.

3. Results and Discussion

3.1. Mapping of the Effect of Operating Conditions on Xylose Recovery in Various Biomasses

The mapping of the pretreatment operating conditions on xylose recovery from various kinds of lignocellulosic biomass: corn fiber, corn cobs, corn stover, palm oil EFB, and sugarcane bagasse, is presented in Figure 1.
Figure 1. Effects of pretreatment operating conditions (CSF) on the xylose recovery from lignocellulosic biomass (a) the xylose recovery is expressed as [mg-xylose/g-dry biomass], (b) the xylose recovery is expressed as the percentage to its maximum/theoretical value, (blue diamonds: corn fiber, red diamonds: corn cob, purple diamonds: corn stover, green squares: oil palm EFB, orange triangles: sugarcane bagasse).

Literature reported moderate pretreatment condition for corn fiber [22]. The calculated CSF value was in the range of 25 – 30 and provided nearly the theoretical recovery of xylose, namely as much as 194 mg/g of corn fiber (Figure 1a-b).
Literatures reported higher pretreatment conditions for corn cob [4,25]. The calculated CSF value was in the range of 7 – 120 and provided the xylose recovery in the range of 70% - 96% of its theoretical value. The highest xylose recovery was obtained at CSF 58.7, giving 58.7% of theoretical xylose recovery which correspond to 302.3 mg/g corn cob (Figure 1a-b). The outlier data of corn cob, CSF 30-90, gave xylose recovery of 45-15% of its theoretical value [6].

Wider data set was obtained on corn stover, the reported pretreatment conditions were in the range of 0 – 230 [5]. At low CSF value that was lower than 100, an increasing trend of xylose recovery was observed along with the increase in the CSF value. Above higher CSF value, a decreasing trend of xylose recovery was observed (Figure 1a-b). The decrease in xylose recovery may be related to the partial xylose degradation at higher operation condition [28]. The maximum xylose recovery was observed in the CSF range of 10-100, giving xylose recovery in the range of 70-95% or 160-240 mg/g corn stover.

A wide range of pretreatment condition was reported for oil palm EFB, ranging from 0 – 267.43 [11,18]. However low xylose recoveries were mostly recorded, in the range of 30-80% or corresponded to 11.7 - 232 mg/g oil palm EFB. The highest xylose recovery of 96.7% was reported by literature, at the application of acid pretreatment method at temperature 120°C, 15 minutes duration, and 12.5% solid loading [19].

A large data set was obtained for sugarcane bagasse [2,26]. An increasing trend of xylose recovery was observed until the CSF of 40 (Figure 1). The maximum xylose recovery (90 – 100%) could be achieved starting at CSF value of 25. Interestingly, a constant trend of high xylose recovery was observed at higher CSF value.

Completely different dataset was obtained from the lignocellulosic biomass that was pretreated by the hydrothermal method. In general, no chemicals, nor acid nor base, is added for the hydrothermal pretreatment method and thus the pH of the system is neutral. Calculating the CSF for this dataset following eq. (2) gives negative values, and thus the pretreatment operation condition is parameterized by SF that is calculated following eq. (1). Similar trend of xylose recovery was observed. The xylose recovery increased along with the increase in severity factor until the maximum value of xylose recovery was achieved. Above that, a decreasing trend of xylose recovery was observed. The maximum xylose recovery was achieved at SF around 4920, giving xylose recovery of 348.87 mg/g hemicellulose which corresponded to xylose recovery of 52% of its theoretical value.
Figure 2. Effects of hydrothermal pretreatment operating conditions (SF) on the xylose recovery from lignocellulosic biomass oil palm EFB. (a) the xylose recovery is expressed as [mg-xylose/g-dry biomass], (b) the xylose recovery is expressed as the percentage to its maximum/theoretical value.

When the trends of the five lignocellulosic biomasses were compared (Figure 1), corn fiber is the biomass that requires the lowest CSF value compared to others, that is in the range of 0 – 30. This shows that corn fiber only needs a moderate pretreatment condition to achieve the maximum xylose recovery. The required mild pretreatment may be related to the presence of starch as well as low lignin content in the corn fiber [22]. Different trend was observed for corn cob, higher CSF value are required, at 0 – 120. Above a certain CSF value a decreasing trend of xylose recovery from corn cob was observed that may be related with the presence of lignin-hemicellulose bonds that cannot be broken by the pretreatment and the further degradation of xylose [28]. The similar trend observed between corn stover and corn cob (Figure 1) may be related to the similarity in the composition of cellulose and lignin in the two biomass. However, the lower xylose recovery that was obtained for the corn stover may be related with the tighter lignin-hemicellulose bonds on the corn stalk and leaf [8]. EFB biomass has data sets in 2 main zones, namely in the CSF value range 0 – 40, xylose recovery 10 – 60% and in the CSF value range 40 – 270, xylose recovery 60 – 100%. This trend indicates that in the first zone there are several lignin-hemicellulose bonds that cannot be removed. However, after passing the CSF value of 40, the bond was able to be released so that the pretreatment was able to effectively break down xylose in the biomass. In addition, what clearly distinguishes EFB from other biomass is the lower xylose recovery under the same pretreatment operating conditions. This may be caused by the higher lignin content of the oil palm EFB [7].

Above discussion indicated that lignin content in the biomass as well as lignin composition may affect the effectivity of the lignocellulosic biomass pretreatment process [24]. The structure of lignin in biomass is expressed by the composition of p-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) contained in lignin biomass [12], and in particular the structure of lignin can be expressed by the ratio of syringyl and guaiacyl [24]. Data on lignin content as well as lignin composition of the studied biomass is presented in Table 1.

Table 1. Lignin and hemicellulose factors of the five lignocellulosic biomass.

| Biomass     | Total hemicellulose | Total lignin | Lignin composition | S/G     | H/G   |
|-------------|---------------------|--------------|--------------------|---------|-------|
| Corn fiber  | 35%[18]             | 8%           | 18.5               | 35.5    | 46.0  |

[18] Data on lignin content as well as lignin composition of the studied biomass is presented in Table 1.
The mapping of the effects of pretreatment operation conditions also shows that maximum xylose recovery can be achieved not only by using a single method. Figure 1 and Figure 2 show that the maximum xylose recovery from oil palm EFB can be achieved either by the weak acid pretreatment method or by the hydrothermal method, despite the different operating conditions required.

3.2. Model Development
An empiric model was developed to determine the effects of pretreatment operating conditions on xylose recovery from lignocellulosic biomass. The data on five types of lignocellulosic biomass: corn fiber, corn cob, corn stover, oil palm EFB, and sugarcane bagasse, as is presented in Figure 1, was used for model development. The operating conditions that we expressed as log CSF, were set as the independent variable, whereas the xylose recovery [mg/g biomass] was set as the dependent variable. Only the increasing trend of xylose recovery was used in the curve fitting feature in the MATLAB, as the decreasing trend of xylose recovery at higher CSF was suspected to be related with further xylose degradation. The best model, however, only gave the R-square value of 0.305 (Table 2). The low R-square value implied that the model could not best represent the data. It indicated that the operating conditions are not the only parameter affecting xylose recovery. Series of model improvements were then conducted to obtain a model that could explain the data better.

### Table 2. Development of Empirical Model on Xylose Recovery.

| Modeling | Equation | Information  |
|----------|----------|-------------|
| Basic model | \( y = a + b \times \log(\text{CSF}) + c \times (\log(\text{CSF}))^2 \) | \( a = -159.6; b = 197.7; c = -20.69 \) | \( R\text{-square} = 0.305 \) (3) |
| incorporating the effect of lignocellulosic biomass composition, i.e. hemicellulose content | \( y = a + b \times \log(\text{CSF}) + c \times (\log(\text{CSF}))^2 \) | \( a = -515.5; b = 564; c = -61.78 \) | \( R\text{-square} = 0.359 \) (4) |
| incorporating the effect of lignocellulosic biomass composition, i.e. lignin content | \( y = a + b \times \log(\text{CSF}) + c \times (\%\text{Lignin}) + d \times (\log(\text{CSF}))^2 + e \times (\%\text{Lignin}) \log(\text{CSF}) + f \times (\%\text{Lignin})^2 \) | \( a = -1694; b = 444; c = 22940; d = -85.3; e = 87970; f = 1254 \) | \( R\text{-square} = 0.5806 \) (5) |
| incorporating the effect lignin composition, i.e. S/G Ratio | \( y = a + b \times \log(\text{CSF}) + c \times \left( \frac{S}{G} \right) + d \times (\log(\text{CSF}))^2 + e \times \left( \frac{S}{G} \right) \log(\text{CSF}) + f \times \left( \frac{S}{G} \right)^2 \) | \( a = -3115; b = 833; c = 2137; d = -78.81; e = 309.4; f = -122.9 \) | \( R\text{-square} = 0.5733 \) (6) |
In the first attempt we changed the unit of xylose recovery to be mg/g of hemicellulose in the biomass (Table 1). This way, we corrected the model by incorporating the effect of lignocellulose biomass composition, in particular the hemicellulose content of the lignocellulose. A better R-square value was obtained, 0.359 (Table 2), which indicated that hemicellulose levels affected the correlation between CSF and xylose recovery. However, the obtained R-square value was still low.

In the second attempt we incorporated the effect of lignin content of the lignocellulosic biomass (Table 1). In this case, two independent variables were introduced to the model, that were the log CSF value and the lignin content of the biomass, whereas the dependent variable was the xylose recovery in mg/g of hemicellulose in the biomass. A better R-square value was obtained, 0.5806 (Table 2). On the other hand, if it was the information on the lignin composition, i.e. the S/G value, that was incorporated into the model instead of the lignin content, a model with R-square value of 0.5733 was obtained (Table 2).

From these two models, it can be seen that the lignin content parameter has a greater correlation than the S/G parameter. However, the S/G equation has a smaller error value than the lignin content. Moreover, the R-square also states the goodness of fit. Therefore, this modeling equation was used to correct xylose recovery by entering the CSF and S/G values for each biomass into each experimental data. The mapping of the corrected xylose recovery results is presented in Figure 3.

![Figure 3. Comparison between data and xylose recovery estimated from model simulation.](image)

The comparison between the data on the five lignocellulosic biomass and the xylose recovery estimated from model simulation is presented in Figure 3. At a similar CSF value, the xylose recovery would be higher on biomass with a higher S/G value. The estimated xylose recovery from corn cob, for example, is always higher than the xylose recovery from oil palm EFB for any value of CSF. However, for corn fiber biomass, the corrected xylose recovery gave negative value. That is, the obtained model can not estimate the xylose recovery correctly. This indicates that still other factor that should be taken into account in the model.

Another correction factor could also be introduced into the model to correct the estimate xylose recovery from lignocellulosic biomass pretreated by hydrothermal pretreatment method. Else, a specific model could also be developed for estimating the xylose recovery directly from the operating condition of the hydrothermal pretreatment method. However, due to data limitation this was not conducted here.
3.3. Model Validation

The developed model (eq. 6) was further validated by using it to estimate the xylose recovery from wheat straw, which data was not used in developing the model. The data on pretreatment operating condition as well as the related xylose recovery from wheat straw was obtained from literature [13].

Figure 4 shows that the developed model could estimate the xylose recovery from wheat straw on any given pretreatment condition quite well. The simulated model gave a similar trend to the data, although the estimated xylose recovery was mostly slightly lower than the data. This may be due to the S/G value of wheat straw was 1.3 [13], lower than the S/G values of the other biomass which were used in the model development. For example, the literature value for the xylose recovery from wheat straw that was obtained from a pretreatment at 170°C and pH 0.4, for 2 minutes, was 20 mg xylose/g of wheat straw. Meanwhile, the model simulation gave an estimation of xylose recovery of 19.4 mg/g wheat straw.

![Figure 4. Model validation, simulation of the developed model to estimate the xylose recovery from wheat straw at any given operating condition (curve: model simulation; literature data [13] symbols: yellow circle).](image)

4. Conclusion

The research has successfully mapped the effects of pretreatment operating conditions on xylose recovery from lignocellulosic materials, in particular on corn fiber, corn cob, corn stover, empty fruit bunches, and sugarcane bagasse. An increase in the pretreatment operating condition leads to an increase in the xylose recovery from lignocellulosic materials. However, above a certain condition, a decreasing trend of xylose recovery that may be related with further xylose degradation can be expected. Besides, the xylose recovery from lignocellulosic materials is also affected by biomass composition, lignin structure of the biomass, and the enzymatic hydrolysis process following to the hydrothermal pretreatment method. An empiric model has been developed to correlate the effect of pretreatment operating condition biomass, composition, lignin structure, and xylose recovery. The model has been successfully validated by using it to estimate the recovery of xylose from other lignocellulosic material, i.e. wheat straw. This model can be further used to estimate the expected xylose recovery from a specific lignocellulosic biomass material at any predesigned pretreatment condition.
Acknowledgments
Authors would like to acknowledge anyone who have contributed in completing the research, either directly or indirectly. In particular, Ms. Diah Meilany is acknowledge for sharing experimental data on EFB pretreatment.

References
[1] Das P K, Das B P and Dash P 2014 Refining Biomass Residues for Sustainable Energy and Bioproducts 2014 Elsevier Inc.
[2] De Paiva J E, Maldonade I R and Scamparini A R P 2009 Revista Brasileira de Engenharia Agrícola e Ambiental 13(1), 75–80.
[3] del Río J C, Lino A G, Colodette J L, Lima C F, Gutiérrez A, Martínez Á T and Rencoret J 2015 Biomass and Bioenergy 81, 322–338.
[4] Dominguez J M, Cao N, Gong C S and Tsao G T 1997 Bioresource Technology 61(1), 85–90.
[5] Feher A, Feher C, Rozbach M and Barta Z 2017 Chemical and Biochemical Engineering Quarterly 31(1), 77–87.
[6] Ghaffar A, Yameen M, Aslam N, Jalal F, Noreen R, Munir B, Mahmood Z, Saleem S, Rafiq N, Falak S, Tahir I M, Noman M, Farooq M U, Qasim S, and Latif F 2017 Omega 11, 97.
[7] Grande P M, Weidener D, Dietrich S, Dama M, Bellof M, Maas R and De Maria P D 2019 ACS Omega 10(1), 201–208.
[8] Guragain Y N, Herrera A I, Vadlani P V and Prakash O 2015 Natural Product Communications 10(1), 201–208.
[9] Harahap B M and Kresnowati M T A P 2018 Biomass Conversion and Biorefinery 8(2), 255–263.
[10] Kabel M A, Bos G, Zeevalking J, Voragen A G J and Schols H A 2007 Bioresource Technology 98(10), 2034–2042.
[11] Li Q, Ng W T, Puah S M, Bhaskar R V, Soh L S, Maclachlan C and Wu J C 2014 Biotechnology and Applied Biochemistry 61(4), 426–431.
[12] Lourenço A and Pereira H 2018 Lignin - Trends and Applications
[13] Lupoi J S, Singh S, Parthasarathi R, Simmons B A and Henry R J 2015 Renewable and Sustainable Energy Reviews 49, 871–906.
[14] Mardawati E, Andoyo R, Syukra K A, Kresnowati M T A P, and Bindar Y 2018 IOP Conference Series: Earth and Environmental Science 141, 012019
[15] Meilany D, Kresnowati M T A P, Setiadi T, and Boopathy R 2020 Applied Sciences (Switzerland), 10(4), 1391.
[16] Modenbach A A 2014 Transactions of the ASABE, 1187–1198.
[17] Nagata T, Lorz H and Widholm J M, 2009 Biotechnology in Agriculture and Forestry February 2015.
[18] Rahman S H A, Choudhury J P, Ahmad A L and Kamaruddin A H 2007 Bioresource Technology 98(3), 554–559.
[19] Rahman S H A, Choudhury J P, and Ahmad A L 2006 Biochemical Engineering Journal 30, 97-103
[20] Romli M, Suprihatin, Indrasti N S and Aryanto A Y Prosiding Seminar Tjipto Utomo Institut Teknologi Nasional Oktober 2010, B71-B7-9.
[21] Ruiz H A, Thomsen M H and Trajano H L 2017 Hydrothermal Processing in Biorefineries: Production of Bioethanol and High Added-Value Compounds of Second and Third Generation Biomass, 1–511.
[22] Saha B C and Bothast R J 1999 Applied Biochemistry and Biotechnology - Part A Enzyme Engineering and Biotechnology 76(2), 65–77.
[23] Scheller H V and Ulvskov P 2010 Annual Review of Plant Biology 61, 263–289.
[24] Studer M H, DeMartini J D, Davis M F, Sykes R W, Davison B, Keller M and Wyman C E 2011
Proceedings of the National Academy of Sciences of the United States of America 108(15), 6300–6305.

[25] Tada K, Horiuchi J I, Kanno T and Kobayashi M 2004 Journal of Bioscience and Bioengineering 98(3), 228–230.

[26] Vargas, Betancur G J and Pereira N 2010 Electronic Journal of Biotechnology 13(3), 1–9.

[27] Venkateswar R L, Goli J K, Gentela J and Koti S 2015 Bioresource Technology 213 (February 2018), 299–310.

[28] Wyman C E, Yang B 2017 Combined Severity Factor for Predicting Sugar Recovery in Acid-Catalyzed Pretreatment Followed by Enzymatic Hydrolysis. 161-180.