

RESEARCH

Open Access



Prediction of phenolic compounds and glucose content from dilute inorganic acid pretreatment of lignocellulosic biomass using artificial neural network modeling

Hongzhen Luo^{1,2*} , Lei Gao¹, Zheng Liu¹, Yongjiang Shi¹, Fang Xie¹, Muhammad Bilal¹, Rongling Yang^{1,3} and Mohammad J. Taherzadeh⁴

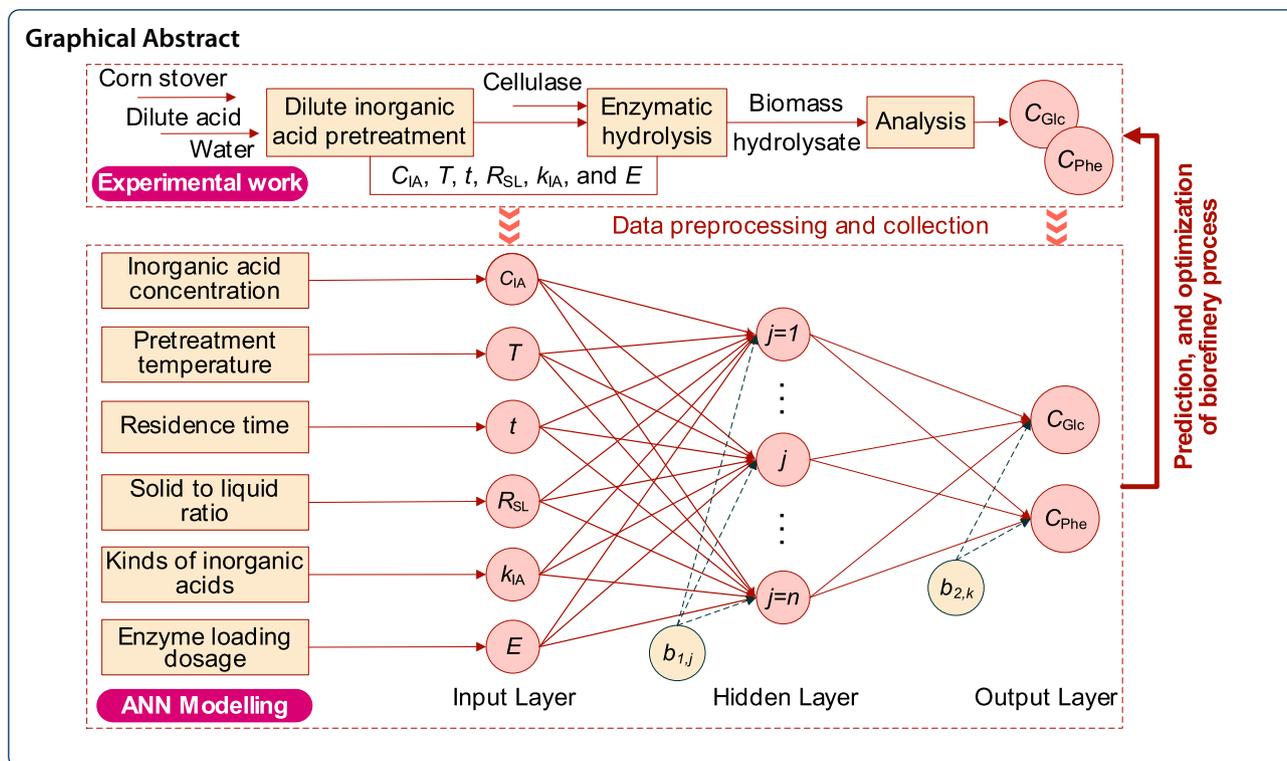
Abstract

Dilute inorganic acids hydrolysis is one of the most promising pretreatment strategies with high recovery of fermentable sugars and low cost for sustainable production of biofuels and chemicals from lignocellulosic biomass. The diverse phenolics derived from lignin degradation during pretreatment are the main inhibitors for enzymatic hydrolysis and fermentation. However, the content features of derived phenolics and produced glucose under different conditions are still unclear due to the highly non-linear characteristic of biomass pretreatment. Here, an artificial neural network (ANN) model was developed for simultaneous prediction of the derived phenolic contents (C_{phe}) and glucose yield (C_{Glc}) in corn stover hydrolysate before microbial fermentation by integrating dilute acid pretreatment and enzymatic hydrolysis. Six processing parameters including inorganic acid concentration (C_{IA}), pretreatment temperature (T), residence time (t), solid-to-liquid ratio (R_{SL}), kinds of inorganic acids (k_{IA}), and enzyme loading dosage (E) were used as input variables. The C_{phe} and C_{Glc} were set as the two output variables. An optimized topology structure of 6–12–2 in the ANN model was determined by comparing root means square errors, which has a better prediction efficiency for C_{phe} ($R^2 = 0.904$) and C_{Glc} ($R^2 = 0.906$). Additionally, the relative importance of six input variables on C_{phe} and C_{Glc} was firstly calculated by the Garson equation with net weight matrixes. The results indicated that C_{IA} had strong effects (22%–23%) on C_{phe} or C_{Glc} , then followed by E and T . In conclusion, the findings provide new insights into the sustainable development and inverse optimization of biorefinery process from ANN modeling perspectives.

Keywords: Lignocellulosic biomass, Dilute acid pretreatment, Enzymatic hydrolysis, Phenolic compounds, Artificial neural network, Modeling

*Correspondence: hzluo@hyit.edu.cn

¹ School of Life Science and Food Engineering, Huaiyin Institute of Technology, 1 Meicheng East Road, Huaian 223003, China
Full list of author information is available at the end of the article



Introduction

Nowadays, the concerns over climate change, especially the increasing greenhouse gases (GHGs) emissions, have necessitated a rethinking of traditional methods for fuels production (Field et al. 2020; Solarte-Toro et al. 2019). According to the statistics, about 25% of the total GHGs emission was contributed by transportation (Keasling et al. 2021). Reducing GHGs emissions and enhancing carbon capture/sequestration (CCS) are global concerns (Ishaq et al. 2021). To this end, one avenue is to produce advanced transportation fuels from atmospheric carbon resources (mainly CO_2) or renewable biomass grown by fixing CO_2 to partially replace fossil resources (Liu et al. 2020).

In the past decades, utilization of lignocellulosic biomass from non-food crops for production of fuels and fine chemicals has gained much attention (Luo et al. 2021a; Rajan et al. 2020) because their net carbon footprint is neutral. Lignin (15–30 wt%), cellulose (30–50 wt%), and hemicellulose (20–35 wt%) are three main components in lignocellulose (Schutyser et al. 2018). The structure of lignocellulose is complex since cellulose and hemicellulose are enwrapped by lignin, and hemicellulose is also interlaced with cellulose fibers (Liu et al. 2021), resulting in the biomass recalcitrance and low enzymatic efficiency for hydrolysis. Depolymerization of lignocellulose to obtain fermentable sugars

(mainly glucose) is the key for production of fuels (Luo et al. 2018). Thus, decreasing biomass recalcitrance and structural complexity via efficient pretreatments is generally required (Liu et al. 2021).

To disrupt the close inter-component association between cellulose, hemicellulose, and lignin, various pretreatment strategies including acid-, alkaline-, ionic liquid-, and organic solvent-based methods have been developed (Hijosa-Valsero et al. 2017; Jönsson and Martín 2016; Xia et al. 2021). Among those pretreatment strategies, dilute inorganic acids hydrolysis is one of the most promising methods with high recovery of fermentable sugars and low cost (Jönsson and Martín 2016), which is beneficial for the production of biofuels and chemicals. Nevertheless, during acid pretreatment of biomass, hemicellulose and lignin are partially solubilized, which result in the degradation of these fragments under an acidic environment (Zhang et al. 2021). Generally, three kinds of inhibitors, including weak organic acids (acetic acid, formic acid, etc.), furan derivatives, and phenolics (phenolic acids, and phenolic aldehydes) are derived during pretreatment, which affect enzymatic and fermentation efficiency (Chen et al. 2020; Yao et al. 2021). Especially, the phenolic compounds derived from lignin degradation during pretreatment characterized by complex structure, diversity, low water solubility, and low hydrophobicity

were reported as the main limiting factor to the industrial biofuel production (Gu et al. 2019).

To counteract the toxic effect of phenolics on the enzymatic and fermentation process, detoxification of lignocellulosic hydrolysates and slurries with overliming, activated carbon, or water washing is widely implemented (Sivagurunathan et al. 2017). However, fermentable sugars were also partially removed, and the treatment of generated wastewater would further deteriorate techno-economic performance. Additionally, the construction of robust strains by elucidating the response mechanism to phenolics could also weaken the inhibitory effect (Kumar et al. 2020; Luo et al. 2021b; Luo et al. 2020). For example, Jiménez-Bonilla et al. (2020) reported that overexpressing efflux pump gene *srpB* from *Pseudomonas putida* is beneficial to improve the tolerance of *Clostridium saccharoperbutylacetonicum* to 1.2 g/L ferulic acid (Jiménez-Bonilla et al. 2020). Although systems metabolic engineering and adaptive evolution strategies could improve the robustness of microbes under biomass-derived inhibitors stress, the derived concentration feature of inhibitors in biomass hydrolysate under pretreatment process should be firstly considered.

Optimization of pretreatment conditions by evaluating glucose hydrolysis efficiency and derived phenolics was carried out using different lignocellulose such as rice straw (Lee et al. 2012), sugarcane bagasse (Lv et al. 2017), etc. The content of lignin-derived phenolics in hydrolysate before fermentation is mainly attributed to pretreatment conditions, such as biomass species, pretreatment temperature, reaction time, solid-to-liquid ratio, etc. (Bhatia et al. 2020; Jönsson and Martín 2016), and also changed during enzymatic hydrolysis (Yao et al. 2021). Implementing numerous experiments for pretreatment of biomass could achieve a sub-optimal result, it would unavoidably increase the operation complexity and time-consuming. Loading of high-cost cellulase also largely improves the total biorefinery cost. Furthermore, the relationship between pretreatment, enzymatic conditions, and derived features of phenolics content was still unclear. Hence, it remains challenging to systematically analyze the effects on derived phenolics and glucose yield concerning both biomass characteristics, pretreatment, and enzymatic conditions.

Development of bioprocess modeling for the non-linear lignocellulosic bioprocessing is an efficient strategy enabling the success of biorefinery and bio-based circular economy (Unrean 2016). Recently, artificial intelligence (AI) technology, mainly machine learning (ML) algorithms, is competent for predicting/confirming relative importance between input and output variables (Li et al. 2021). The effectiveness of ML methods for predicting pyrolytic gas yield and compositions was verified (Tang

et al. 2021), which could benefit to better understand biomass pyrolysis and syngas upgrading. An artificial neural network (ANN) model with a multilayer architecture (3-15-1) was optimized and predicted the biogas production curve from cattle under mesophilic and thermophilic conditions (Ghatak and Ghatak 2018). An ANN model was built to predict sugar yields of pretreated rice straw during hydrolysis by considering three factors of biomass loadings, particle size, and reaction time (Vani et al. 2015). Recently, Moodley et al. (2019) found that sugar yield from sugarcane leaf waste was sensitive to the alkali and salt concentrations by establishing two ANN tools in inorganic salt pretreatment process (Moodley et al. 2019). The above reports clearly show that ANN can be trained with experimental data to generate efficient models of non-linear multivariate processes. To the best of our knowledge, prediction of lignin-derived phenolics content and glucose concentration from inorganic acid pretreatment of biomass with advanced modeling technology was not reported thus far.

Focusing on above-mentioned issues, in this study, we aim to develop an ANN model for elucidating the derived feature of phenolics from corn stover by integrative investigating typical three inorganic acids (HCl, H₂SO₄, and H₃PO₄) pretreatment and enzymatic hydrolysis processes. Furthermore, the relative importance of pretreatment conditions (i.e., input variables) on phenolic and glucose concentrations (i.e., output variables) was also first elucidated by considering the neural net weights in the developed ANN model. The results would provide new insights into the biorefinery process for biofuels production.

Materials and methods

Materials and chemicals

The corn stover was collected from Lianyungang City, China. It was firstly cut and sieved to a particle size of ~0.4 mm. The fine corn stover was then dried in an oven (GZX-9140MBE, Shanghai Boxun Medical Biological Instrument Corp., China) at 60 °C for 12 h to remove the moisture and stored in plastic bags at 4 °C. Three kinds of inorganic acids, including dibasic acid (hydrochloric acid, 37 wt%), binary acid (sulfuric acid, 98 wt%), and ternary acid (phosphoric acid, 85 wt%) were used as pretreatment agents for corn stover depolymerization. The inorganic acids were purchased from Sinopharm Chemical Reagent Co., Ltd. All of the chemicals were used as received without other specified purification. A commercial cellulase Cellic CTec2 (enzyme blend, SAE0020-50 mL solution) was obtained from Sigma-Aldrich (St. Louis, MO, USA), and used for hydrolyzing the pretreated corn stover to obtain glucose.

Table 1 The design of operating conditions to perform the dilute acid pretreatment of lignocellulosic biomass for the development of ANN model

Lignocellulose	Kinds of acids (k_{IA}) ^a	Acid concentration (C_{IA} , mol/L)	Pretreatment temperature (T , °C)	Residence time (t , min)	Solid to liquid ratio (R_{SL})	Enzyme loading dosage (E , FPU/g)
Corn stover	Hydrochloric acid (HCl)	0.05–0.6	120–200	20–60	10–15%	10–20
	Sulfuric acid (H ₂ SO ₄)	0.05–0.6	120–200	20–60	10–15%	10–20
	Phosphoric acid (H ₃ PO ₄)	0.05–0.6	120–200	20–60	10–15%	10–20

The pretreatment experiments with specified conditions were all carried out with three replicates. The raw data used for ANN development were the means and the detailed means \pm SD values are shown in Table S1

Three kinds of inorganic acid including dibasic acid (HCl), binary acid (H₂SO₄), and ternary acid (H₃PO₄) were used

^a The k value for HCl, H₂SO₄ and H₃PO₄ was set as 1, 2, and 3, respectively, which was used as one input parameter for the development of ANN model

Dilute inorganic acid pretreatment and enzymatic hydrolysis

To systematically investigate the effects of pretreatment conditions on the derived phenolic compounds and glucose hydrolysis yield, six key parameters including inorganic acid concentration (C_{IA}), pretreatment temperature (T), residence time (t), solid-to-liquid ratio (R_{SL}), kinds of inorganic acids (k_{IA}), and enzyme loading dosage (E) were considered in this study. The acid pretreatment process of corn stover was carried out in a 250-mL vertical reactor (TGYF-B, Gongyi Yuhua Instrument Co., Ltd., China) with an electrically magnetic stirrer and a temperature controller. Firstly, the dried corn stover, inorganic acid, and 150 mL water were added into the reactor simultaneously. The experimental parameter ranges for corn stover pretreatment, and enzymatic hydrolysis were carefully designed and also summarized in Table 1, containing the raw data for training, validation, and testing the following ANN model. After acid hydrolysis, the pH of the pretreated mixture was regulated to 5.0 by 8 mol/L NaOH solution. The overall effects of biomass-derived phenolic compounds on enzymatic hydrolysis were also considered due to the interaction of phenolics with cellulose. Thus, the pretreated mixture was not filtrated, and directly used for enzymatic hydrolysis by adding cellulase with 10–20 FPU/g corn stover (Table 1). The reaction was conducted at 50 °C in a water bath at 150 rpm. After 72 h enzymatic hydrolysis, the pretreated mixture was firstly boiled at 95–100 °C for 5 min to terminate the reaction. Then, the liquid fraction of the mixture (i.e., hydrolysate) was separated by vacuum filtration for determining phenolics and glucose contents.

Development of ANN model

Selection of input variables and experimental data

Different kinds of biomass (corn stover, rice straw, switchgrass, sugarcane straw, etc.) have different ratios of cellulose, hemicellulose, and lignin, which result in various features of the derived phenolics and glucose

yield even under the same pretreatment and enzymatic hydrolysis conditions (Pratto et al. 2020; Solarte-Toro et al. 2019). Among these lignocellulosic biomasses, corn stover is the largest crop residue in China (Yang et al. 2020b); thus, it was selected as the model biomass for the pretreatment experiments in this study.

The formation of the biomass-derived phenolic compounds content and glucose hydrolysis yield in hydrolysate is a complex and non-linear process, which is difficult to directly predict the derived features with traditional constructive mathematical models. Thus, we tried to use ANN modeling to predict derived phenolic compounds content and glucose yield after pretreatment and enzymatic hydrolysis processes. As described in above section, six key parameters of C_{IA} , T , t , R_{SL} , k_{IA} , and E (Table 1) were considered as the input variables for the development of ANN model. The output variables were glucose concentration (C_{Glc}) and phenolic content (C_{Phe}) in biomass hydrolysate after 72 h enzymatic hydrolysis. To better perform the ANN model, 77 runs experiments were carried out according to the design of input variables shown in Table 1.

Data preprocessing, and the topology of ANN model

Figure 1A shows the step-by-step scheme of ANN model development to predict C_{Phe} and C_{Glc} from corn stover under different operation conditions. The ANN model was developed by using Matlab R2019a (The MathWorks, Inc., USA). It has one multiple layer neural network with interconnected neurons arranged in three layers of input, hidden, and output layers (Fig. 1B). The ANN model proposed in present study consists in: (1) two parameters of C_{Glc} and C_{Phe} were considered the output variables; (2) six variables of C_{IA} , T , t , R_{SL} , k_{IA} , and E in one input layer were fully connected to the hidden layer; (3) one hidden layer had n neurons; and (4) bias of $b_{1,j}$ (the bias of inputs, $j = 1, 2, 3, \dots, n$) and $b_{2,k}$ (the bias of output layer, $k = 1, 2$) were used for training the network.

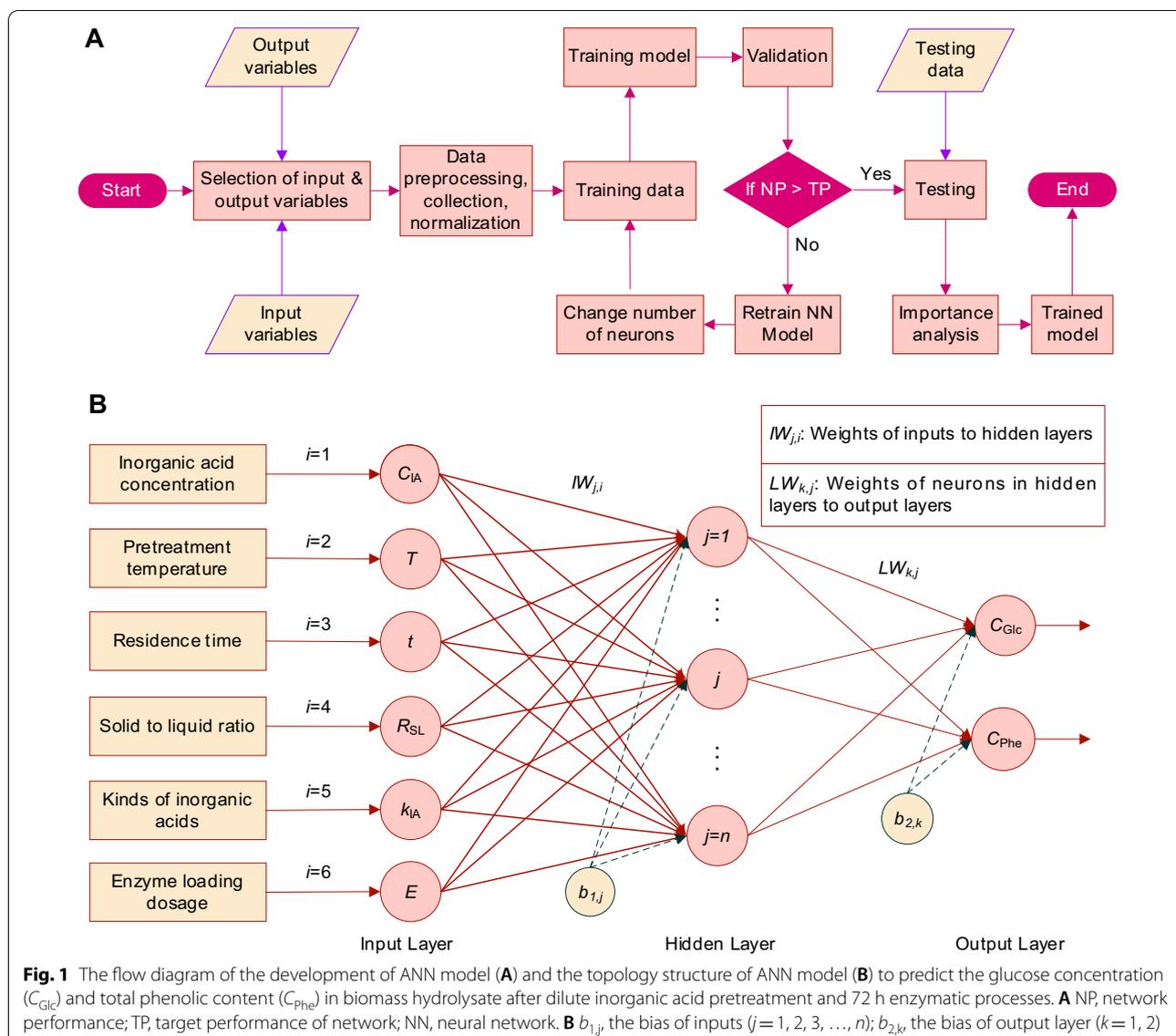


Fig. 1 The flow diagram of the development of ANN model (A) and the topology structure of ANN model (B) to predict the glucose concentration (C_{Glc}) and total phenolic content (C_{Phe}) in biomass hydrolysate after dilute inorganic acid pretreatment and 72 h enzymatic processes. A NP, network performance; TP, target performance of network; NN, neural network. B $b_{1,j}$, the bias of inputs ($j=1, 2, 3, \dots, n$); $b_{2,k}$, the bias of output layer ($k=1, 2$)

Firstly, the whole datasets were split as training, validation, and testing groups with a ratio of 75:15:10. The input dataset is normalized in the range of $[-1, 1]$ before training the network to obtain an accurate model. The inbuilt ‘mapminmax’ function is used for the normalization of experimental data, which is equivalent to Eq. (1):

$$V' = \frac{V - V_{\min}}{V_{\max} - V_{\min}} (V'_{\max} - V'_{\min}) + V'_{\min}, \quad (1)$$

where V' , V , V_{\max} , V_{\min} , V'_{\max} , and V'_{\min} represented the new value, the original value, the original maximum limit, original minimum limit, the new maximum limit (i.e., 1), and the new minimum limit (i.e., -1), respectively.

The ANN model was trained by the Adam optimizer, with a learning rate of 0.001, and training batches of size

2. The neurons (n) in the hidden layer were determined by an empirical equation Eq. (2) (Yang et al. 2020a). The root means square error (RMSE) obtained from different neurons and iterations were used to evaluate the accuracy of model predictions, which was calculated by Eq. (3). Based on RMSE results, an optimized ANN model with better performance was developed to predict C_{Glc} and C_{Phe} (Fig. 1A):

$$n = \sqrt{i + k} + \alpha, \quad (2)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{h=1}^m (y_{pre}^{(h)} - \hat{y}_{exp}^{(h)})^2}, \quad (3)$$

where n is the number of neurons in hidden layer; i is the number of input variables; k is the number of output variables; α is a constant range of 1–10; $y_{\text{pre}}^{(h)}$ is the predicted output value of C_{Phe} ; $\hat{y}_{\text{exp}}^{(h)}$ is the experimental value of the output variable of C_{Phe} ; and m is the number of the samples for training, validation, or testing of ANN models.

Analysis of relative importance of input variables

The parameters of $IW_{j,i}$, $LW_{k,j}$, $b_{1,j}$, and $b_{2,k}$ in the developed ANN model could be used to simulate the output variables (C_{Glc} , C_{Phe}). In addition, to evaluate the relative importance of the input variables on the two output variables, the process was based on the neural net weight matrixes ($IW_{j,i}$ and $LW_{k,j}$, Fig. 1B) and Garson equation (Puig-Arnavat et al. 2013; Sunphorka et al. 2017). Garson equation was based on the partitioning of connection weights in the network. The numerator presents the total of absolute products of weights for each input ($i = 1, 2, \dots, 6$) while the denominator represents the total of the absolute values of all weights feeding into the hidden layer ($j = 1, 2, \dots, n$). The Garson equation is presented in Eq. (4) for adapting the ANN topology:

$$I_i = \frac{\sum_{j=1}^n \left(\left(|IW_{j,i}| \div \sum_{i=1}^6 |IW_{j,i}| \right) \times |LW_{k,j}| \right)}{\sum_{i=1}^6 \left\{ \sum_{j=1}^n \left(\left(|IW_{j,i}| \div \sum_{i=1}^6 |IW_{j,i}| \right) \times |LW_{k,j}| \right) \right\}} \times 100\%, \quad (4)$$

where I_i is the relative importance of the i th input variable on output variables of C_{Glc} and C_{Phe} ; $IW_{j,i}$ is the neural net weight to j th neuron of the hidden layer from i th input variable; and $LW_{k,j}$ is the weight to k th output variable from j th neuron of the hidden layer, respectively.

Analytical methods

The glucose concentration in corn stover hydrolysate (C_{Glc}) was determined by a biosensor analyzer (S-10, Sie-man Technology, China) (Luo et al. 2019). Determination of total phenolic content (C_{Phe}) in the hydrolysate was based on the Folin–Ciocalteu assay with gallic acid as the standard (Xu et al. 2021) with some modifications. Briefly, 0.4 mL sample was mixed with 2.6 mL water and 0.5 mL Folin–Ciocalteu reagent (1.0 mol/L, Sinopharm Chemical Reagent Co., Ltd., Shanghai). After 5 min, 5.0 mL water and 1.5 mL Na_2CO_3 solution (20%, w/v) was added simultaneously. Then, the mixture oscillated under a dark environment at 40 °C for 1 h. Finally, the absorbance of the reaction mixture was analyzed at 760 nm by a UV–Vis spectrophotometer (UV-2100, Unico Instrument Co., Ltd., China). As a result, the C_{Phe} (g/L) was calculated by Eq. (5):

$$C_{\text{Phe}} = \frac{a \times A_{760}}{V_s} \times N, \quad (5)$$

where a is the linear coefficient of standard curve; A_{760} is the absorbance of reaction mixture at 760 nm; V_s is the volume of reaction mixture; and N is the dilution ratio, respectively.

Statistical analysis

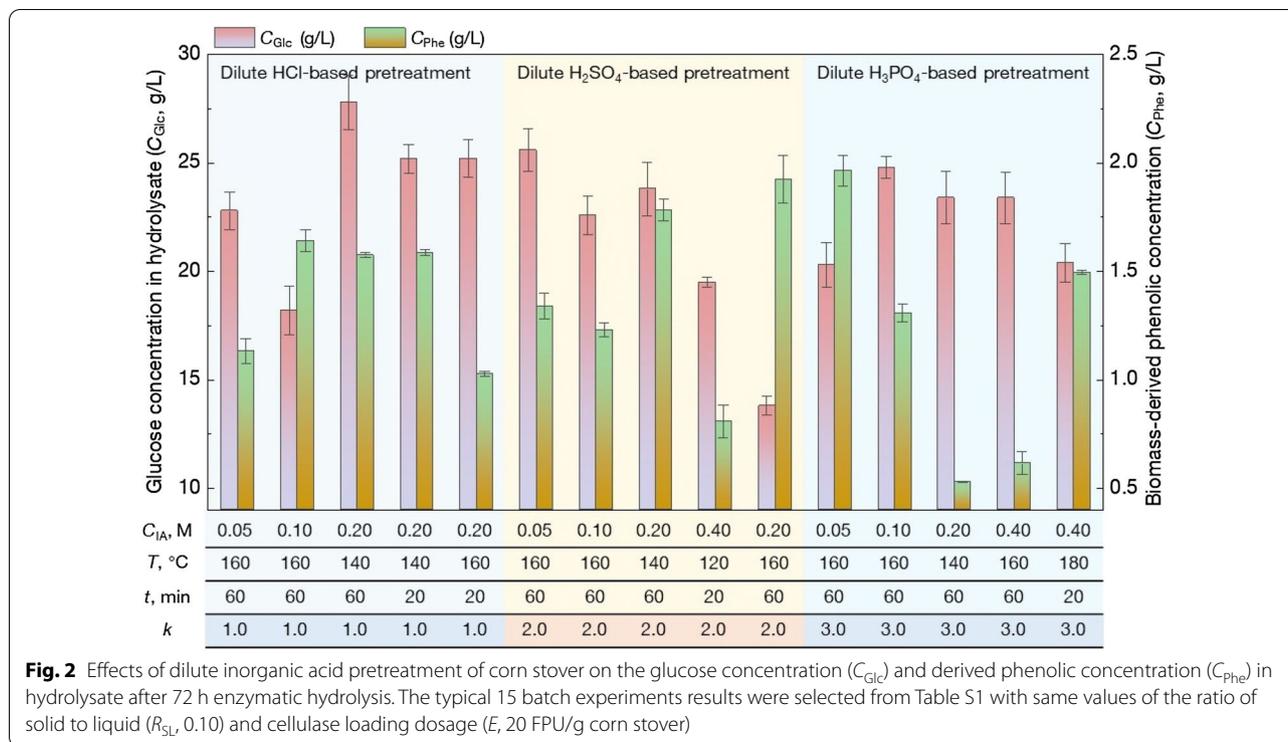
The experimental data of C_{Phe} and C_{Glc} for development of ANN model are represented as the mean \pm standard deviation (SD) of three independent experiments. Significant differences were confirmed with a two-tailed Student's t -test aided by Microsoft Excel 2016.

Results

Effects of inorganic acid pretreatment/enzymatic hydrolysis on the content of derived phenolics and glucose

For efficient production of biofuels and fine chemicals from lignocellulosic biomass via microbial fermentation, an optimized pretreatment process featured with a high glucose yield from feedstock and a low derived concentration of inhibitors is crucial (Liu et al. 2021). Inorganic acid-based pretreatment is applied to efficiently solubilize hemicellulose from lignocellulose and it also improves the cellulose digestibility (Jönsson and Martín 2016; Zabed et al. 2016). Therefore, dilute inorganic acids hydrolysis is widely used in biorefinery process. In this study, focusing on dilute inorganic acid pretreatment of corn stover, the overall effects of pretreatment conditions on phenolics and glucose concentration after 72 h enzymatic hydrolysis were studied, and the results are shown in Fig. 2. Three typical inorganic acids, including HCl (dibasic acid), H_2SO_4 (binary acid), and H_3PO_4 (ternary acid) were used as the pretreatment reagent. It should be noted, in these cases, the solid-to-liquid ratio (R_{SL}), and cellulase loading dosage (E) was kept at 10% and 20 FPU/g corn stover, respectively.

When pretreating corn stover with 0.05 mol/L of HCl under 160 °C for 60 min, the derived phenolic compounds concentration in hydrolysate (C_{Phe}) was 1.13 g/L, and glucose content (C_{Glc}) reached 22.8 g/L. Under the same condition of H_2SO_4 pretreatment, C_{Glc} were increased by 12.3% (25.6 g/L), and C_{Phe} was also elevated to 1.34 g/L with significant differences ($p < 0.05$) to the HCl pretreatment process. Whereas, in the case of 0.05 mol/L H_3PO_4 pretreatment with 160 °C for 60 min, the glucose yield was only 20.3 g/L with a higher level of C_{Phe} (1.97 g/L, Fig. 2). The results indicated that cellulase might tolerate a low concentration of C_{Phe} . For the HCl pretreatment of corn stover, C_{Phe} was increased to 1.64 g/L when elevating acid concentration from 0.05 to 0.1 mol/L, but the phenomenon was not found in



H₂SO₄ and H₃PO₄ pretreatment processes. In addition, the effects of different R_{SL} and E on C_{Glc} and C_{Phe} were also studied and the detailed experimental data are also shown in Additional file 1: Table S1. Since glucose and derived phenolic contents are mainly attributed to complex pretreatment conditions and enzymatic process with multivariate non-linear features (Huang et al. 2007), it is still challenging to speculate the changing patterns of C_{Glc} and C_{Phe} by only implementing numerous experiments. Hence, it is necessary to explore advanced methods for elucidating the relative importance of operational conditions on C_{Glc} or C_{Phe} for further optimizing the lignocellulose pretreatment process.

Optimization and determination of key parameters in ANN model

As shown in Fig. 1B, the multilayer ANN model for the prediction of C_{Glc} and C_{Phe} consisted of 6 neurons in one input layer, one hidden layer, and one output layer. Since the number of neurons in the hidden layer is a key parameter in ANN model, and trial-error approach was applied to ensure a relatively fast and good convergence of RMSE (Puig-Arnavat et al. 2013). The obtained results of 11 ANN models under different neuron numbers (n) with 500 iterations are exhibited in Table 2. Here, an empirical equation Eq. (2) was used to determine the range of neuron numbers (from 3 to 13) in the hidden layer (Yang et al. 2020a). As shown in Table 2, when the neurons in

Table 2 Computational results of RMSE during the training and validation processes with different neurons of hidden layer in ANN models. RMSE is the root means square error calculated by Eq. (3)

Neurons in hidden layer, n	RMSE for C_{Glc} after 500 iterations		RMSE for C_{Phe} after 500 iterations	
	Training dataset	Validation dataset	Training dataset	Validation dataset
$n=3$	7.16	7.12	0.54	0.41
$n=4$	9.35	8.51	0.87	0.42
$n=5$	7.92	6.06	0.56	0.18
$n=6$	6.71	6.41	1.37	1.05
$n=7$	6.44	5.47	0.46	0.22
$n=8$	7.53	5.84	0.46	0.46
$n=9$	6.41	5.03	0.42	0.40
$n=10$	6.40	5.20	0.43	0.28
$n=11$	6.57	4.97	0.53	0.25
$n=12$	6.38	4.73	0.43	0.24
$n=13$	6.50	5.46	0.43	0.32

the hidden layer was 3, RMSE for C_{Glc} reached 7.16 in the training dataset and 7.12 in the validation dataset. For the RMSE of C_{Phe} , a range of 0.41–0.54 was obtained in the case of 3 neurons. When changing the neurons from 3 to 13 in the hidden layer, RMSE varied due to the different parameters while developing ANN models. By

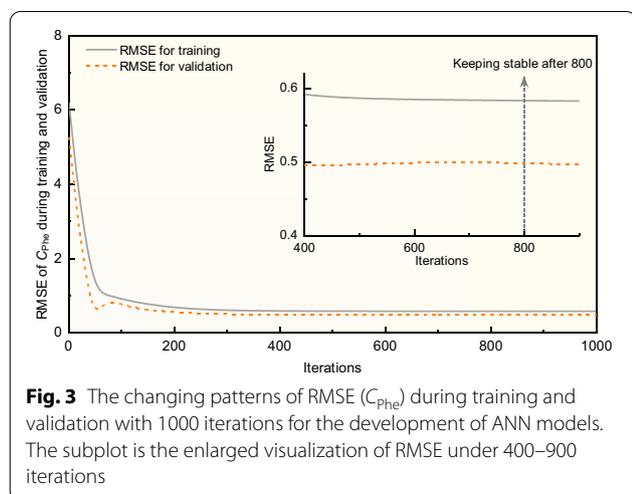


Fig. 3 The changing patterns of RMSE (C_{Phe}) during training and validation with 1000 iterations for the development of ANN models. The subplot is the enlarged visualization of RMSE under 400–900 iterations

combination analysis of the RMSE for C_{Glc} and C_{Phe} , the network performance of ANN model with 12 neurons was better which showed relative lower values of RMSE (Table 2). In addition, the effects of iterations on RMSE for C_{Phe} when training and validation of the ANN models were also investigated. The RMSE changing patterns in Fig. 3 indicated that increasing iterations from 400 to 800 could effectively decrease the RMSE of C_{Phe} and then kept stable after 800 iterations. Therefore, an optimized iteration of 800 with the best network structure of 6-12-2 was considered.

Training and testing of the ANN model

The best ANN structure was achieved after numerical experiments using training and validation datasets with a single hidden layer consisted of $n = 12$ neurons (Fig. 1B; Table 2). During the training process with 57 batches dataset (each including 6 input variables and 2 output variables, Additional file 1: Table S1), the predicted values of C_{Glc} and C_{Phe} are exhibited in Fig. 4. For the predicted values of C_{Glc} during training process, the RMSE was 5.77 (Fig. 4A), and the corresponding result for C_{Phe} reached a lower value of 0.44 (Fig. 4B). In addition, the optimized weights ($IW_{j,i}$ and $LW_{k,j}$) and biases ($b_{1,j}$ and $b_{2,k}$) of the proposed ANN model are listed in Table 3.

To verify the effectiveness of the proposed ANN model, a testing dataset with suitable ranges of input/output variables was also used to predict C_{Glc} and C_{Phe} under different dilute inorganic acid pretreatment conditions. The fitting relationships between the predicted and experimental values of C_{Glc} and C_{Phe} are plotted in Fig. 5. The slope of the fitting curve and the correlation coefficient of R^2 value are the two key parameters for accurately evaluation of the proposed ANN model. In other words,

a better fitting relationship between predicted values obtained from the developed model and experimental values generally featured with the slope and R^2 nearing to 1.0. As shown in Fig. 5, the diagonal of the plot ($Y = X$) was displayed with a slope of 1.0 for clearly comparison. It is indicated that the R^2 of C_{Glc} fitting curve was 0.906 under the range of 7.5–25 g/L, with a slope of 0.86. In addition, the experimental values of C_{Phe} were located at 0.8–3.0 g/L. The ANN model obtained the fitting curve with a slope of 0.82 and a R^2 of 0.904 for prediction of C_{Phe} . Based on the fitting performance shown in Fig. 5, it is concluded that ANN modeling is an efficient tool for predicting C_{Glc} and C_{Phe} simultaneously from the non-linearity and complexity of the input–output system containing corn stover pretreatment/enzymatic processes.

Relative importance of input parameters on output variables

For the pretreatment of lignocellulosic biomass, elucidating and understanding the influence of pretreatment or enzymatic conditions on glucose hydrolysis yield and inhibitors formation are beneficial for guiding the biorefinery process (Bhatia et al. 2020). Although numerous studies have investigated the effects of pretreatment and enzymatic methods on glucose yield and inhibitors formation (Hassan et al. 2018; Jönsson and Martín 2016), the quantitative relationship properties between operation variables and those derived contents in biomass hydrolysate before fermentation are still unclear.

Focusing on this concern, the relative importance (I_i , $i = 1, 2 \dots 6$) of the six input variables on the two output variables of C_{Glc} and C_{Phe} were analyzed (Fig. 6). The I_i was calculated by the Garson equation Eq. (4) with the weight matrixes of $IW_{j,i}$ and $LW_{k,j}$ in the developed ANN model, which are listed in Table 3. As shown in Fig. 6A, the six input variables strongly influenced C_{Glc} with the range of 12–23%. The five parameters included in the pretreatment process (C_{IA} , T , t , R_{SL} , and k_{IA}) represent up to 79% importance, and enzyme dosage (E) accounts for 21% importance on C_{Glc} . Interestingly, the highest importance ($I_1 = 23\%$) on C_{Glc} is the concentration of inorganic acid (C_{IA}), which is even higher than that of enzyme dosage (21%). The results revealed that an efficient pretreatment strategy is beneficial for glucose hydrolysis, which is mainly attributed to the improved accessibility of cellulose during pretreatment process (Siqueira et al. 2017; Xu et al. 2016). The relative importance of T , R_{SL} , and k_{IA} on C_{Glc} is around 14%–16%, while in this case, the I_3 (t) is the lowest index (Fig. 6A). In addition, the relative importance of input variables on C_{Phe} is plotted in Fig. 6B. Similarly, inorganic acid concentration (C_{IA}) also has the strongest importance of 22% on C_{Phe} . It is indicated that a severe acidic environment with a higher concentration of

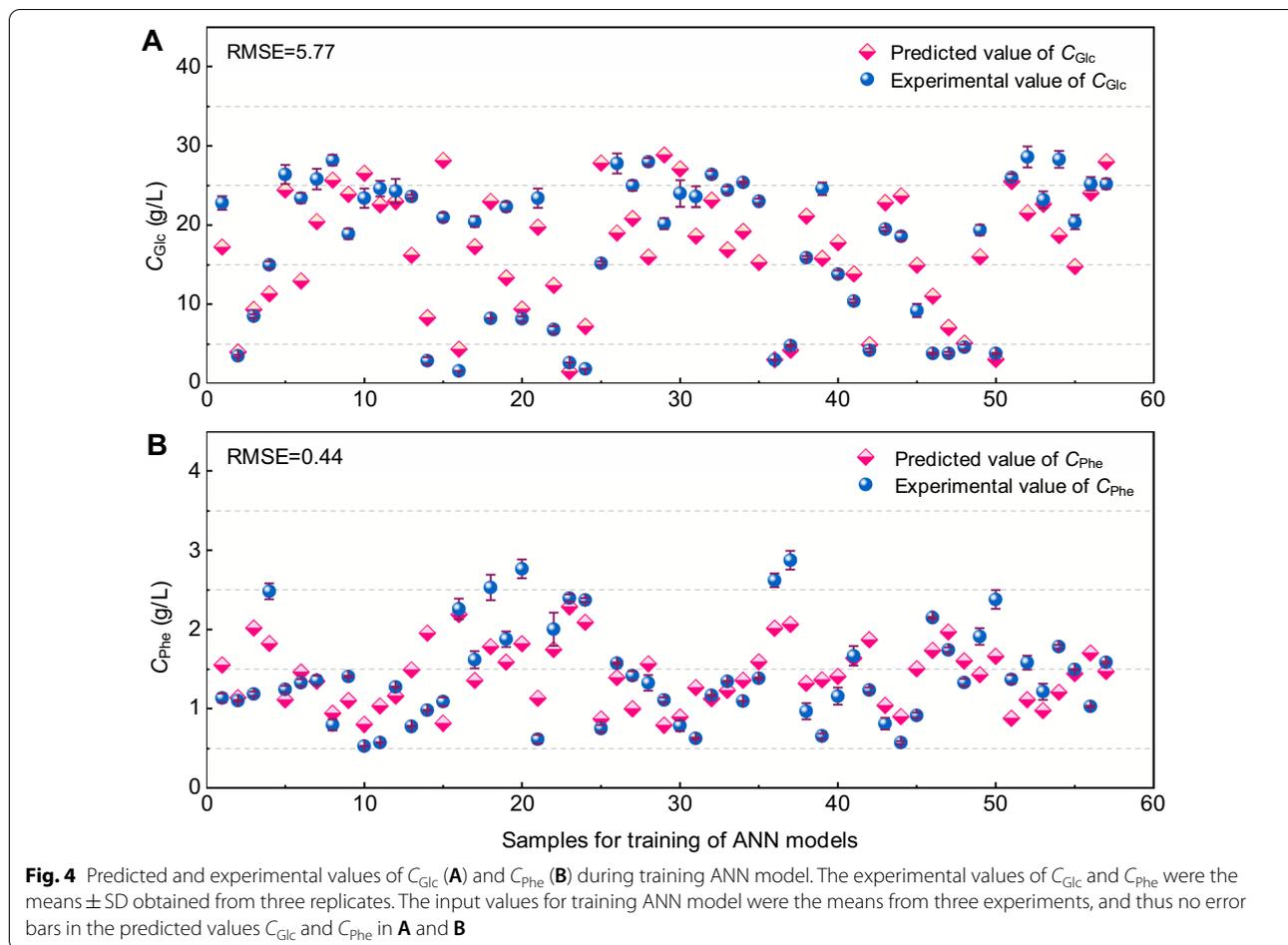


Table 3 Weights and biases of the hidden and output layers used in the developed ANN model for the prediction of C_{Glc} and C_{Phe}

Node, j	Weights and biases of the hidden layer							Weights and biases of the output layer		
	$IW_{j,1}$	$IW_{j,2}$	$IW_{j,3}$	$IW_{j,4}$	$IW_{j,5}$	$IW_{j,6}$	$b_{1,j}$	$LW_{1,j}$	$LW_{2,j}$	$b_{2,k}$
1	0.4211	0.5806	0.1236	0.2909	0.1691	0.5445	-0.0095	0.0448	0.4378	-0.0381
2	0.3931	0.1473	0.2595	0.4280	0.2980	0.4272	-0.0393	0.3972	0.4253	-0.0051
3	0.3056	0.1740	0.3713	0.1522	0.1229	0.1367	-0.0287	0.0356	0.2299	
4	0.5184	0.0507	0.2491	0.0508	0.3027	0.4435	-0.0379	0.6842	0.1607	
5	0.3112	0.4807	0.7085	0.2731	0.3305	0.6146	0.0394	0.2159	0.0281	
6	0.1854	0.5232	0.0776	0.4782	0.3439	0.1783	-0.0306	0.1249	0.4419	
7	0.0370	0.2334	0.0454	0.3650	0.2354	0.4857	0.0372	0.3608	0.0639	
8	0.2985	0.2512	0.3173	0.1907	0.1834	0.5006	0.0361	0.5057	0.4299	
9	0.3070	0.4097	0.0005	0.5172	0.3769	0.2365	-0.0184	0.0140	0.3469	
10	0.3136	0.2505	0.1438	0.0805	0.0552	0.1178	-0.0320	0.2159	0.6152	
11	0.6273	0.0770	0.3091	0.3422	0.3044	0.3722	0.0394	0.6076	0.3147	
12	0.5133	0.5409	0.0001	0.4964	0.2515	0.1423	0.0371	0.5611	0.2426	

j is the neuron node in the hidden layer (from $j=1$ to 12). IW_{ji} ($i=1, 2, 3, 4, 5,$ and 6) is the neural net weight to j th neuron of the hidden layer from i th input variable; and LW_{kj} ($k=1$ and 2) is the neural net weight to k th output variable from j th neuron of the hidden layer. $b_{1,j}$ is the bias of input variables, and $b_{2,k}$ is the bias of output layer

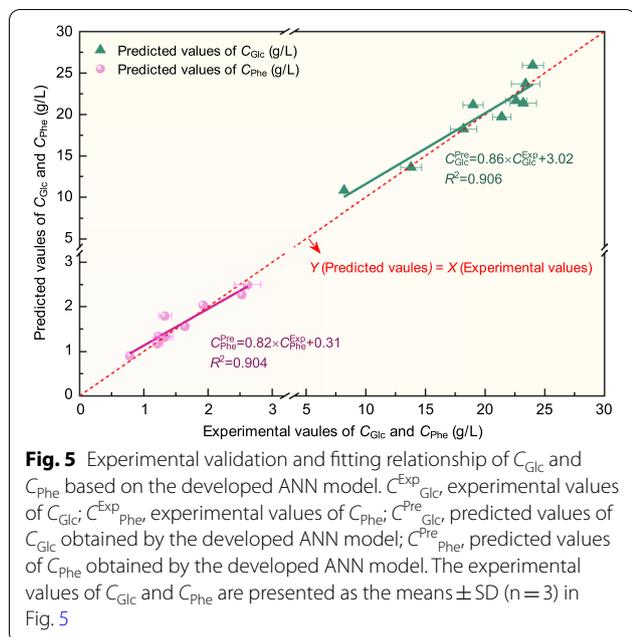


Fig. 5 Experimental validation and fitting relationship of C_{Glc} and C_{Phe} based on the developed ANN model. C_{Glc}^{Exp} , experimental values of C_{Glc} ; C_{Phe}^{Exp} , experimental values of C_{Phe} ; C_{Glc}^{Pre} , predicted values of C_{Glc} obtained by the developed ANN model; C_{Phe}^{Pre} , predicted values of C_{Phe} obtained by the developed ANN model. The experimental values of C_{Glc} and C_{Phe} are presented as the means \pm SD ($n = 3$) in Fig. 5

H^+ could improve the efficiency of lignin deconstruction coupling with phenolics formation (He et al. 2020). The other four variables in the pretreatment process occupied 60% importance on C_{Phe} (i.e., T for 19%, t for 11%, R_{SL} for 17%, and k_{IA} for 13%, Fig. 6B). It should be noted that enzyme dosage (E) still kept higher importance of 18% on C_{Phe} (Fig. 6B), with a relatively lower importance of 21% on C_{Glc} (Fig. 6A). The obtained results of the relative importance of enzyme dosage on C_{Phe} reflected that the derived phenolic contents would change in biomass hydrolysate during enzymatic hydrolysis.

Discussion

Optimization of biorefinery process can lead to highly efficient production of biofuels from renewable resources such as lignocellulosic biomass. Traditional optimization method called “one variable at time” (OVAT) is time-consuming and requires a large number of experiments. To circumvent the limitation, experimental models are used to elucidate the relationship between operating parameters and final outcomes. Recently, various models such as response surface methodology (RSM) and ANN were reported (Das et al. 2015; Fernandes et al. 2020; Sewsynker-Sukai and Gueguim Kana 2018). The obstacle of RSM is the limitation to the hypothesis of quadratic correlation between conditions because it assumes the second-order polynomial equation (Fernandes et al. 2020). Therefore, ANN modeling was used to build efficient models to predict C_{Phe} and C_{Glc} .

Some ANN models were recently developed to assess the biomass pretreatment for the production of biofuels and fine chemicals (Moodley et al. 2019; Sunphorka et al. 2017; Vani et al. 2015). Compared with those models, the proposed ANN model in this study possessed three advantages. Firstly, the derived feature of the diversity of phenolic compound content from dilute inorganic acid pretreatment of lignocellulosic biomass, which is one of the most promising approaches to industrial implementation, was clarified when using corn stover as the feedstock for the first time. Secondly, the dilute inorganic acid pretreatment and enzymatic hydrolysis processes were considered from a systematic perspective. It is because C_{Phe} and C_{Glc} in biomass hydrolysate are the key factors affecting the overall fermentation performance. Lastly, the relationship of relative importance between operation

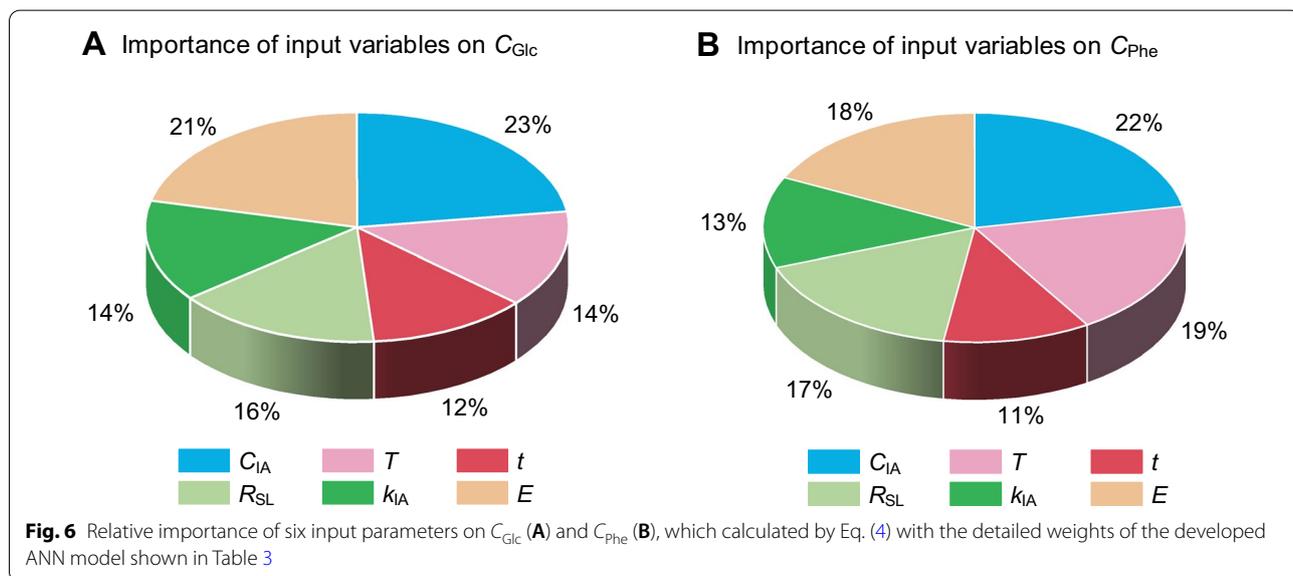


Fig. 6 Relative importance of six input parameters on C_{Glc} (A) and C_{Phe} (B), which calculated by Eq. (4) with the detailed weights of the developed ANN model shown in Table 3

variables and output variables (C_{Phe} and C_{Glc}) was clearly elucidated by the weight matrixes in the developed ANN model. It should be noted that although ANN models could not accurately calculate/predict results beyond the range of operational parameters in training/validation datasets, it can still provide an estimation of parameters in an uncharted workspace by catching the trends during the training process (Rashid et al. 2021). Collection of previously reported data for the development of advanced ANN models would be an efficient strategy. Therefore, it is concluded that ANN modeling is a powerful tool for predicting key parameters in some crucial multivariate non-linear bioprocesses. The relative importance analysis also provides new insights into the biochemical process assessment and optimization.

Although the phenolic compounds were directly formed by lignin degradation during pretreatment process, some studies found that the content (C_{Phe}) still changes during the enzymatic hydrolysis process. It is mainly attributed to the interactions effect between cellulase and lignin-derived phenolics (Yao et al. 2021; Zhao et al. 2021). In addition, the water-soluble lignin-derived phenolics were adsorbed by cellulase and inhibit the enzymatic efficiency (Yuan et al. 2021). Those findings could explain the result that the enzyme dosage has a high importance of 18% on C_{Phe} (Fig. 6B).

Although dilute inorganic acid pretreatment was selected to elucidate the derived feature of phenolic content from corn stover, some other indispensable factors should also be considered to further improve the sustainability of biorefinery process. Different pretreatment methods have various solubilization abilities of lignin, cellulose, and hemicellulose. For the removal efficiency of lignin, the alkaline-based hydrolysis is generally higher than that of dilute acid-based methods (Zabed et al. 2016). If the aim is only to investigate lignin removal efficiency from lignocellulosic biomass, alkaline-based pretreatment might be a better choice. In addition, the operation parameters including the particle size and biomass species also influenced the fermentable sugars and derived phenolic contents (Vani et al. 2015), which should be explored in the future.

Conclusions

An artificial neural network (ANN) model was developed to simultaneously predict the derived feature of phenolic compounds content (C_{Phe}) and glucose yield (C_{Glc}) in biomass hydrolysate from dilute inorganic acid pretreatment and enzymatic hydrolysis. Five pretreatment and enzyme dosage parameters were used as the input variables in the optimized ANN model, which has one hidden layer with 12 neurons. Results indicated that the developed ANN model has a good fitting performance ($R^2 > 0.90$) for the

prediction of C_{Phe} and C_{Glc} . The relative importance of six variables on C_{Phe} and C_{Glc} was also calculated to provide new insights for optimizing biorefinery to produce biofuels.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40643-021-00488-x>.

Additional file 1: Table S1. The input data for development of ANN model to predict C_{Glc} and C_{Phe} in biomass hydrolysate after dilute inorganic acid pretreatment and enzymatic hydrolysis.

Acknowledgements

Not applicable.

Authors' contributions

HL: supervision, conceptualization, methodology, funding acquisition, writing—reviewing and editing; LG, ZL, and YS: investigation, software, methodology, writing—original draft; FX: data curation, visualization; MB: investigation, writing—review and editing; RY: investigation, visualization; MJT: investigation, writing—review and editing.

Funding

This work is supported by the National Natural Science Foundation of China (21808075), the Natural Science Foundation of Jiangsu Province (BK20170459), and the Science and Technology Innovation Project of Huaiyin Institute of Technology (HGYK202106).

Availability of data and materials

The data supporting the conclusions of this article are included in the main manuscript.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All of the authors have read and approved to submit it to *Bioresources and Bioprocessing*.

Competing interests

The authors declare that they have no competing interests.

Author details

¹School of Life Science and Food Engineering, Huaiyin Institute of Technology, 1 Meicheng East Road, Huaian 223003, China. ²Jiangsu Provincial Engineering Laboratory for Biomass Conversion and Process Integration, Huaiyin Institute of Technology, Huaian 223003, China. ³Faculty of Applied Technology, Huaiyin Institute of Technology, Huaian 223003, China. ⁴Swedish Centre for Resource Recovery, University of Borås, 50190 Borås, Sweden.

Received: 13 October 2021 Accepted: 12 December 2021

Published online: 19 December 2021

References

- Bhatia SK, Jagtap SS, Bedekar AA, Bhatia RK, Patel AK, Pant D, Rajesh Banu J, Rao CV, Kim Y-G, Yang Y-H (2020) Recent developments in pretreatment technologies on lignocellulosic biomass: Effect of key parameters, technological improvements, and challenges. *Bioresour Technol* 300:122724
- Chen X, Zhai R, Li Y, Yuan X, Liu Z-H, Jin M (2020) Understanding the structural characteristics of water-soluble phenolic compounds from four

- pretreatments of corn stover and their inhibitory effects on enzymatic hydrolysis and fermentation. *Biotechnol Biofuels* 13:44
- Das S, Bhattacharya A, Haldar S, Ganguly A, Gu S, Ting YP, Chatterjee PK (2015) Optimization of enzymatic saccharification of water hyacinth biomass for bio-ethanol: comparison between artificial neural network and response surface methodology. *Sustain Mater Techno* 3:17–28
- Fernandes CD, Nascimento VRS, Meneses DB, Vilar DS, Torres NH, Leite MS, Vega Baudrit JR, Bilal M, Iqbal HMN, Bharagava RN, Egues SM, Romanholo Ferreira LF (2020) Fungal biosynthesis of lignin-modifying enzymes from pulp wash and *Luffa cylindrica* for azo dye RB5 biodecolorization using modeling by response surface methodology and artificial neural network. *J Hazard Mater* 399:123094
- Field JL, Richard TL, Smithwick EAH, Cai H, Laser MS, LeBauer DS, Long SP, Paus-tian K, Qin Z, Sheehan JJ, Smith P, Wang MQ, Lynd LR (2020) Robust paths to net greenhouse gas mitigation and negative emissions via advanced biofuels. *Proc Natl Acad Sci USA* 117:21968–21977
- Ghatak MD, Ghatak A (2018) Artificial neural network model to predict behavior of biogas production curve from mixed lignocellulosic co-substrates. *Fuel* 232:178–189
- Gu H, Zhu Y, Peng Y, Liang X, Liu X, Shao L, Xu Y, Xu Z, Liu R, Li J (2019) Physiological mechanism of improved tolerance of *Saccharomyces cerevisiae* to lignin-derived phenolic acids in lignocellulosic ethanol fermentation by short-term adaptation. *Biotechnol Biofuels* 12:268
- Hassan SS, Williams GA, Jaiswal AK (2018) Emerging technologies for the pre-treatment of lignocellulosic biomass. *Bioresour Technol* 262:310–318
- He J, Huang C, Lai C, Huang C, Li M, Pu Y, Ragauskas AJ, Yong Q (2020) The effect of lignin degradation products on the generation of pseudo-lignin during dilute acid pretreatment. *Ind Crop Prod* 146:112205
- Hijosa-Valsero M, Paniagua-García AI, Diez-Antolínez R (2017) Biobutanol production from apple pomace: the importance of pretreatment methods on the fermentability of lignocellulosic agro-food wastes. *Appl Microbiol Biotechnol* 101:8041–8052
- Huang J, Mei LH, Xia J (2007) Application of artificial neural network coupling particle swarm optimization algorithm to biocatalytic production of GABA. *Biotechnol Bioeng* 96:924–931
- Ishaq H, Ali U, Sher F, Anus M, Imran M (2021) Process analysis of improved process modifications for ammonia-based post-combustion CO₂ capture. *J Environ Chem Eng* 9:104928
- Jiménez-Bonilla P, Zhang J, Wang Y, Bliersch D, de Bashan L-E, Guo L, Wang Y (2020) Enhancing the tolerance of *Clostridium saccharoperbutylacetoni-cum* to lignocellulosic-biomass-derived inhibitors for efficient biobutanol production by overexpressing efflux pumps genes from *Pseudomonas putida*. *Bioresour Technol* 312:123532
- Jönsson LJ, Martín C (2016) Pretreatment of lignocellulose: formation of inhibitory by-products and strategies for minimizing their effects. *Bioresour Technol* 199:103–112
- Keasling J, Garcia Martin H, Lee TS, Mukhopadhyay A, Singer SW, Sundstrom E (2021) Microbial production of advanced biofuels. *Nat Rev Microbiol* 19:701–715
- Kumar V, Yadav SK, Kumar J, Ahluwalia V (2020) A critical review on current strategies and trends employed for removal of inhibitors and toxic materials generated during biomass pretreatment. *Bioresour Technol* 299:122633
- Lee K-M, Kalyani D, Tiwari MK, Kim T-S, Dhiman SS, Lee J-K, Kim I-W (2012) Enhanced enzymatic hydrolysis of rice straw by removal of phenolic compounds using a novel laccase from yeast *Yarrowia lipolytica*. *Bioresour Technol* 123:636–645
- Li J, Zhang W, Liu T, Yang L, Li H, Peng H, Jiang S, Wang X, Leng L (2021) Machine learning aided bio-oil production with high energy recovery and low nitrogen content from hydrothermal liquefaction of biomass with experiment verification. *Chem Eng J* 425:130649
- Liu Z, Wang K, Chen Y, Tan T, Nielsen J (2020) Third-generation biorefineries as the means to produce fuels and chemicals from CO₂. *Nat Catal* 3:274–288
- Liu Y, Cruz-Morales P, Zargar A, Belcher MS, Pang B, Englund E, Dan Q, Yin K, Keasling JD (2021) Biofuels for a sustainable future. *Cell* 184:1636–1647
- Luo H, Yang R, Zhao Y, Wang Z, Liu Z, Huang M, Zeng Q (2018) Recent advances and strategies in process and strain engineering for the production of butyric acid by microbial fermentation. *Bioresour Technol* 253:343–354
- Luo H, Zheng P, Xie F, Yang R, Liu L, Han S, Zhao Y, Bilal M (2019) Co-production of solvents and organic acids in butanol fermentation by *Clostridium acetobutylicum* in the presence of lignin-derived phenolics. *RSC Adv* 9:6919–6927
- Luo H, Zheng P, Bilal M, Xie F, Zeng Q, Zhu C, Yang R, Wang Z (2020) Efficient bio-butanol production from lignocellulosic waste by elucidating the mechanisms of *Clostridium acetobutylicum* response to phenolic inhibitors. *Sci Total Environ* 710:136399
- Luo H, Liu Z, Xie F, Bilal M, Liu L, Yang R, Wang Z (2021a) Microbial production of gamma-aminobutyric acid: applications, state-of-the-art achievements, and future perspectives. *Crit Rev Biotechnol* 41:491–512
- Luo H, Liu Z, Xie F, Bilal M, Peng F (2021b) Lignocellulosic biomass to biobutanol: Toxic effects and response mechanism of the combined stress of lignin-derived phenolic acids and phenolic aldehydes to *Clostridium acetobutylicum*. *Ind Crop Prod* 170:113722
- Lv X, Xiong C, Li S, Chen X, Xiao W, Zhang D, Li J, Gong Y, Lin J, Liu Z (2017) Vacuum-assisted alkaline pretreatment as an innovative approach for enhancing fermentable sugar yield and decreasing inhibitor production of sugarcane bagasse. *Bioresour Technol* 239:402–411
- Moodley P, Rorke DCS, Gueguim Kana EB (2019) Development of artificial neural network tools for predicting sugar yields from inorganic salt-based pretreatment of lignocellulosic biomass. *Bioresour Technol* 273:682–686
- Pratto B, Chandgude V, de Sousa R, Cruz AJG, Bankar S (2020) Biobutanol production from sugarcane straw: Defining optimal biomass loading for improved ABE fermentation. *Ind Crop Prod* 148:112265
- Puig-Arnaut M, Hernández JA, Bruno JC, Coronas A (2013) Artificial neural network models for biomass gasification in fluidized bed gasifiers. *Biomass Bioenergy* 49:279–289
- Rajan K, Elder T, Abdoulmoumine N, Carrier DJ, Labbé N (2020) Understanding the in situ state of lignocellulosic biomass during ionic liquids-based engineering of renewable materials and chemicals. *Green Chem* 22:6748–6766
- Rashid T, Taqvi SAA, Sher F, Rubab S, Thanabalan M, Bilal M, ul Islam B, (2021) Enhanced lignin extraction and optimisation from oil palm biomass using neural network modelling. *Fuel* 293:120485
- Schutysse W, Renders T, Van den Bosch S, Koelewijn SF, Beckham GT, Sels BF (2018) Chemicals from lignin: an interplay of lignocellulose fractionation, depolymerisation, and upgrading. *Chem Soc Rev* 47:852–908
- Sewsynker-Sukai Y, Gueguim Kana EB (2018) Microwave-assisted alkalic salt pretreatment of corn cob wastes: process optimization for improved sugar recovery. *Ind Crop Prod* 125:284–292
- Siqueira G, Arantes V, Saddler JN, Ferraz A, Milagres AMF (2017) Limitation of cellulose accessibility and unproductive binding of cellulases by pre-treated sugarcane bagasse lignin. *Biotechnol Biofuels* 10:176
- Sivagurunathan P, Kumar G, Mudhoo A, Rene ER, Saratale GD, Kobayashi T, Xu K, Kim S-H, Kim D-H (2017) Fermentative hydrogen production using lignocellulose biomass: an overview of pre-treatment methods, inhibitor effects and detoxification experiences. *Renew Sust Energ Rev* 77:28–42
- Solarte-Toro JC, Romero-García JM, Martínez-Patiño JC, Ruiz-Ramos E, Castro-Galiano E, Cardona-Alzate CA (2019) Acid pretreatment of lignocellulosic biomass for energy vectors production: a review focused on operational conditions and techno-economic assessment for bioethanol production. *Renew Sust Energ Rev* 107:587–601
- Sunphorka S, Chalermisinsuan B, Piumsomboon P (2017) Artificial neural network model for the prediction of kinetic parameters of biomass pyrolysis from its constituents. *Fuel* 193:142–158
- Tang Q, Chen Y, Yang H, Liu M, Xiao H, Wang S, Chen H, Raza Naqvi S (2021) Machine learning prediction of pyrolytic gas yield and compositions with feature reduction methods: effects of pyrolysis conditions and biomass characteristics. *Bioresour Technol* 339:125581
- Unrean P (2016) Bioprocess modelling for the design and optimization of lignocellulosic biomass fermentation. *Bioresour Bioprocess* 3:1
- Vani S, Sukumaran RK, Savithri S (2015) Prediction of sugar yields during hydrolysis of lignocellulosic biomass using artificial neural network modeling. *Bioresour Technol* 188:128–135
- Xia Q, Chen C, Yao Y, Li J, He S, Zhou Y, Li T, Pan X, Yao Y, Hu L (2021) A strong, biodegradable and recyclable lignocellulosic bioplastic. *Nat Sustain* 4:627–635
- Xu GC, Ding JC, Han RZ, Dong JJ, Ni Y (2016) Enhancing cellulose accessibility of corn stover by deep eutectic solvent pretreatment for butanol fermentation. *Bioresour Technol* 203:364–369

- Xu L, Zhu L, Dai Y, Gao S, Wang Q, Wang X, Chen X (2021) Impact of yeast fermentation on nutritional and biological properties of defatted adlay (*Coix lachryma-jobi* L.). *LWT Food Sci Technol* 137:110396
- Yang J, Huang Y, Xu HY, Gu DY, Xu F, Tang JT, Fang C, Yang Y (2020a) Optimization of fungi co-fermentation for improving anthraquinone contents and antioxidant activity using artificial neural networks. *Food Chem* 313:126138
- Yang X, Han D, Zhao Y, Li R, Wu Y (2020b) Environmental evaluation of a distributed-centralized biomass pyrolysis system: a case study in Shandong, China. *Sci Total Environ* 716:136915
- Yao L, Yang H, Yoo CG, Chen C, Meng X, Dai J, Yang C, Yu J, Ragauskas AJ, Chen X (2021) A mechanistic study of cellulase adsorption onto lignin. *Green Chem* 23:333–339
- Yuan Y, Jiang B, Chen H, Wu W, Wu S, Jin Y, Xiao H (2021) Recent advances in understanding the effects of lignin structural characteristics on enzymatic hydrolysis. *Biotechnol Biofuels* 14:205
- Zabed H, Sahu JN, Boyce AN, Faruq G (2016) Fuel ethanol production from lignocellulosic biomass: an overview on feedstocks and technological approaches. *Renew Sust Energy Rev* 66:751–774
- Zhang H, Han L, Dong H (2021) An insight to pretreatment, enzyme adsorption and enzymatic hydrolysis of lignocellulosic biomass: experimental and modeling studies. *Renew Sust Energy Rev* 140:110758
- Zhao X, Meng X, Ragauskas AJ, Lai C, Ling Z, Huang C, Yong Q (2021) Unlocking the secret of lignin-enzyme interactions: recent advances in developing state-of-the-art analytical techniques. *Biotechnol Adv*. <https://doi.org/10.1016/j.biotechadv.2021.107830>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
