



OPEN

Stochastic modelling of cellulose hydrolysis with Gauss and Weibull distributed transition probabilities

Joseph Mcgreg Duru, Oana Cristina Pârvulescu✉, Tănase Dobre & Cristian Eugen Răducanu

Two Markov-type stochastic models were developed to describe the kinetics of acid hydrolysis of cellulose. One of them involved a Gauss (normal) distribution of probabilities of chemical bond breaking, the other a Weibull distribution. It was considered that the random breaking of cellulose was based on the cleavage of a parent macromolecule into two descendants. Model equations and kinetics of acid hydrolysis of cellulose consisting of 10 and 100 units of cellobiose were presented. The effects of acid concentration and temperature on the kinetics of hydrolysis process were taken into account. The results obtained applying both stochastic models were in a reasonable agreement with those obtained using a deterministic kinetic model. These stochastic models can accurately describe the kinetics of acid hydrolysis and cover the drawbacks of some deterministic kinetic models, e.g., large number of model equations and parameters, modification of parameter values by changing the process conditions.

Nomenclature

c_a	Acid concentration (%)
c_C	Concentration of cellulose in the reaction medium (g _C /L)
c_{CB}	Concentration of cellobiose in the reaction medium (g _{CB} /L)
c_{C0}	Initial concentration of cellulose in the reaction medium (g _C /L)
c_{E1}	Concentration of enzymes E1 in the reaction medium (FPU/L) (FPU-filter paper units)
c_G	Concentration of glucose in the reaction medium (g _G /L)
c_{HMF}	Concentration of hydroxymethylfurfural in the reaction medium (g _{HMF} /L)
c_i^s	Concentration of i hydrolysable species within the molecular mass class M_i^s (g _i /L)
CV	Coefficient of variation
E_A	Activation energy (J/mol)
f	Probability density function
k_j	Kinetic constant for j reaction (s ⁻¹)
K_L	Langmuir constant for sorption equilibrium (FPU/L)
k_{maxj}	Maximum value of k_j at full saturation of the substrate with enzymes E1 (s ⁻¹)
M_i^s	Molecular mass class (range) of s polysaccharide species (g _i /mol)
PD	Polymerization degree
P_{ji}^s	Transition probability of s polysaccharide species from M_i^s to M_j^s
p^s	Probability of breaking of s polysaccharide species
P^s	Transition probability matrix of s polysaccharide species
R	Ideal gas constant (8.314 J/mol/K)
$RMSE$	Root mean square error (g _i /L)
T	Absolute temperature (K)
v_R	Reaction rate (g/L/s)
X	Random variable
x	Value of X
α_{ji}^s	Transition probability frequency of s polysaccharide species (s ⁻¹)
β	Shape parameter
λ	Scale parameter
μ	Mean of x values
σ	Standard deviation

Chemical and Biochemical Engineering Department, University POLITEHNICA of Bucharest, 1-6 Gheorghe Polizu, 011061 Bucharest, Romania. ✉email: oana.purvulescu@yahoo.com

σ^2	Variance
τ	Time (s)
ω_i^s	Mass fraction of s polysaccharide species within M_i^s (g/g _C)

Subscripts

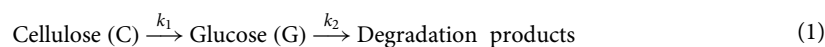
C	Cellulose
CB	Cellobiose
G	Glucose
GS	Gauss
HMF	Hydroxymethylfurfural
ref	Reference
W	Weibull

Superscript

s	Polysaccharide species ($s = 1, 2, \dots, S$)
-----	---

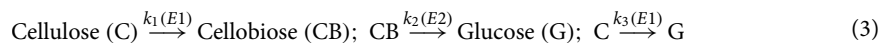
Thermal or (bio)chemical degradation of biopolymers, e.g., cellulose, hemicellulose, starch, lignin, other low and medium molecular mass polysaccharides, has been studied extensively for the last twenty years^{1–10}. Chemical or enzymatic depolymerization of cellulose or starch occurs in aqueous suspension and is known as hydrolysis. Despite of the fact that the cellulose or starch hydrolysis is a heterogeneous (bio)chemical process, it is often analyzed using formal kinetics.

The first models describing dilute acid hydrolysis of cellulose were based on pseudo-homogeneous irreversible first-order reactions, e.g., those presented in scheme (1), where k_1 and k_2 are kinetic constants^{11–15}. According to scheme (1), time (τ) variation of glucose concentration (c_G) can be expressed by Eq. (2), where c_{C0} is the initial cellulose concentration.



$$c_G(\tau) = c_{C0} \left(\frac{k_1}{k_1 - k_2} \right) \left(e^{-k_1\tau} - e^{-k_2\tau} \right) \quad (2)$$

In the case of enzymatic hydrolysis, three basic processes, which are shown in scheme (3), where k_1 , k_2 , and k_3 are kinetic constants, whereas $E1$ and $E2$ are specific enzymes, can be assumed. Three types of enzymes are commonly considered in the biochemical hydrolysis of cellulose, i.e., endocellulases, exocellulases, and cellobiases (β -glucosidases). The mode of action of these enzymes is as follows: (i) endocellulases break cellulose chains into shorter polysaccharide chains; (ii) exocellulases attack from the ends of these shorter polysaccharide chains resulting in oligosaccharides, predominantly cellobiose; (iii) cellobiases cleave the cellobiose into glucose. In the presence of endo/exocellulases ($E1$) and β -glucosidases ($E2$), the processes presented in scheme (3) can be described by simple first-order kinetics expressed by Eqs. (4)–(6), where c_C and c_{CB} are concentrations of cellulose and cellobiose in the reaction medium, v_{R1} , v_{R2} , and v_{R3} are reaction rates. Reaction rate constants for the first and third process depend on the adsorption kinetics of enzymes on cellulose chains^{16–18}. Assuming a rapid adsorption of enzymes $E1$, the process occurs near to equilibrium. Accordingly, considering a Langmuir adsorption isotherm, k_1 and k_3 are given by Eqs. (7) and (8), where c_{E1} is the concentration of enzymes $E1$ in the reaction medium, K_L represents the Langmuir constant for sorption equilibrium, $k_{\max 1}$ and $k_{\max 3}$ are the maximum values of k_1 and k_3 at full saturation of the substrate with enzymes $E1$ for the first and third process in the scheme (3)¹⁹. To determine the reaction rate constant for the second process (transformation of cellobiose into glucose), the activity and concentration of enzymes $E2$ in liquid phase as well as the fact that there is a classic enzyme catalysis described by Michaelis–Menten equation must be taken into account¹⁹. Moreover, the effect of process temperature, reaction inhibition by product and substrate should be considered^{18–23}.



$$v_{R1} = k_1 c_C \quad (4)$$

$$v_{R2} = k_2 c_{CB} \quad (5)$$

$$v_{R3} = k_3 c_C \quad (6)$$

$$k_1 = \frac{k_{\max 1} c_{E1}}{K_L + c_{E1}} \quad (7)$$

$$k_3 = \frac{k_{\max 3} c_{E1}}{K_L + c_{E1}} \quad (8)$$

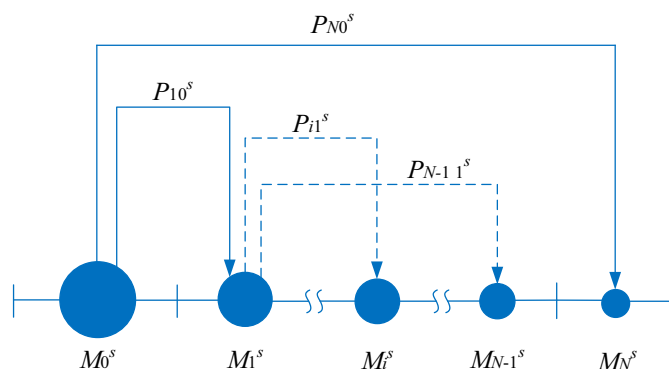


Figure 1. Stochastic model principle in terms of hydrolysate molecular mass intervals and transition probabilities for a polysaccharide species ($s = 1, 2 \dots S$).

The decomposition of (bio)polymers leads to a large number of components involved in complex schemes of parallel and consecutive reactions. For example, in the thermal degradation of halogenated polymers, i.e., poly(chloroprene) and poly(vinyl chloride), kinetic mechanisms taking into account 38–40 species and pseudo-components (molecules and radicals) involved in 190–250 chemical reactions were presented^{24,25}. It is very difficult in this case to apply mathematical models based on formal kinetics. Models assuming random scission of linear polymer chains²⁶ and chain-end scission²⁷ can be more appropriate. Mechanisms based on systematic breakage of polymer chains explain quite well the hydrolysis process of cellulose^{28,29}. Markov chain was used to describe the enzymatic and acid hydrolysis of cellulose³⁰. Markov-type stochastic models are commonly applied in all engineering fields where elementary processes with random evolution appear^{31–33}.

Stochastic models of cellulose hydrolysis, which are based on two basic phenomena, i.e., breaking or non-breaking of polymer chain, could be effective and more realistic approaches. On the one hand, the concentration of cellulose-derived products and frequency distribution of molecular chain length during hydrolysis can be easily determined by a stochastic model, which can provide a better understanding of the mechanisms involved in the process. On the other hand, working with large transition probability matrices, imposed by the initial length of cellulose molecular chain or by the existence of several initial chain lengths, can be computationally quite complex.

Assuming that the highest probability of bond breaking is in the middle of the polymer chain, the transition probabilities can be expressed using a Gauss (normal distribution). A Weibull distribution can be applied if the breaking is more likely to occur at the ends of the polymer chain. For example, in the biochemical hydrolysis of cellulose, it can be assumed that endocellulases act in the middle of the cellulose chain, whereas exocellulases attack from the ends of polymer chains.

This paper aimed at studying the acid hydrolysis of cellulose. A Markov-type stochastic model was developed, assuming that the probabilities of chemical bond breaking followed either a Gauss (normal) distribution or a Weibull distribution.

Methods

The stochastic model focuses on partial hydrolysates obtained by acid hydrolysis of cellulose, which are numerically characterized by their molecular masses. The basic assumptions of the model are as follows: (i) cellulose subjected to acid hydrolysis contains a finite number of hydrolysable polysaccharide species (cellulose fragments with different polymerization degrees) with known molecular masses (M_0^s , $s = 1, 2 \dots S$) and concentrations; (ii) random breaking of a polysaccharide into several macromolecules is based on the cleavage of a parent macromolecule into two descendants; (iii) breaking of the polysaccharide chain into different fragments can occur with different probabilities; (iv) the continuous distribution of molecular masses of hydrolysates can be divided into discrete intervals with a molecular mass corresponding to each interval; (v) a hydrolysate within a molecular mass range (class) is divided into two macromolecules within lower molecular mass ranges; (vi) the process of molecular fragmentation by hydrolysis is a homogeneous Markov process; (vii) hydrolysates cannot participate in any coupling reactions.

Stochastic model principle of hydrolysis of a polysaccharide species ($s = 1, 2 \dots S$) is shown in Fig. 1. The breaking of a macromolecule within M_0^s molecular mass class into one within M_1^s molecular mass class and another within M_N^s molecular mass class is represented with continuous line. The process evolution, i.e., the macromolecule within M_1^s molecular mass class splits into one within M_i^s molecular mass class and another within M_{N-1}^s molecular mass class, appears with dotted line. Transition probability is denoted by P_{ab}^s , where the subscripts a and b refer to the molecular mass ranges of hydrolysates after and before breaking, respectively.

The mass balance for s species parent macromolecules within the molecular mass class M_k^s which break at time τ into descendants within the molecular mass classes M_i^s and M_m^s is given by Eq. (9), where m_k^s represents the mass of parent macromolecules, whereas m_i^s and m_m^s are the masses of descendants. Moreover, if a parent macromolecule breaks into two descendants, the probability of its breaking by hydrolysis (p_k^s) is equal to the probability of birth of each descendant (P_{ik}^s and P_{mk}^s), as shown by Eq. (10).

$$m_k^s = m_l^s + m_m^s \quad (9)$$

$$p_k^s = P_{lk}^s = P_{mk}^s \quad (10)$$

For a short time interval ($\Delta\tau$), breaking probability of macromolecules within the molecular mass class M_i^s , p_i^s ($i=0, 1 \dots N$), and breaking missing probability associated with this class, P_{ii}^s , are defined by Eqs. (11) and (12). Assuming that the hydrolysis of s species is a Markov homogeneous stochastic process, the transition probability from the molecular mass class M_i^s to M_j^s , P_{ji}^s ($i, j=0, 1 \dots N, j \neq i$), is defined by Eq. (13), where α_{ji}^s is the transition probability frequency and $\Delta\tau$ the time interval. The transition probability in mass units, $P_{ji}^{s,m}$, is given by Eq. (14) depending on molecular mass classes and transition frequencies. Unsteady state species mass balance led to Eq. (15), expressing time variation of mass fraction of s species within the molecular mass range M_i^s , ω_i^s .

$$p_i^s = \sum_{\substack{j=1 \\ j \neq i}}^N P_{ji}^s \quad (11)$$

$$P_{ii}^s = 1 - p_i^s = 1 - \sum_{\substack{j=1 \\ j \neq i}}^N P_{ji}^s \quad (12)$$

$$P_{ji}^s = \alpha_{ji}^s \Delta\tau \quad (13)$$

$$P_{ji}^{s,m} = \frac{M_j^s \alpha_{ji}^s}{\sum_{n=1}^N M_n^s \alpha_{ni}^s} \quad (14)$$

$$\frac{d\omega_i^s}{d\tau} = \sum_{\substack{h=0 \\ h \neq i}}^N P_{ih}^{s,m} \omega_h^s - \omega_i^s \sum_{\substack{j=1 \\ j \neq i}}^N P_{ji}^{s,m} \quad (15)$$

Current state of mass fractions of hydrolysable species within the molecular mass classes $M_0^1, M_1^1 \dots M_{10}^1$ (Fig. 2) is given by the probability vector described by Eq. (16), where τ_n , defined by Eq. (17), represents the current time and n is the number of time sequences taken into account in the evolution of hydrolysis process. Initial state of probability vector, which is expressed by Eq. (18), corresponds to hydrolysis starting, when there is only the polysaccharide within the molecular mass class M_0^1 .

Results and discussions

Acid hydrolysis of cellulose consisting of 10 units of cellobiose. The stochastic model was applied to describe a hypothetical hydrolysis of a polysaccharide species ($s=1$) consisting of 10 units of cellobiose. As shown in Fig. 2, 10 uniformly distributed molecular mass classes ($M_0^1=3258$ g/mol, $M_1^1=2934$ g/mol, $M_2^1=2610$ g/mol, ..., $M_8^1=666$ g/mol, and $M_9^1=342$ g/mol) were taken into account. Moreover, cellobiose disaccharide ($M_9^1=342$ g/mol) was hydrolyzed resulting in 2 molecules of glucose ($M_{10}^1=180$ g/mol). Schema in Fig. 2 is based on data reported in the related literature^{28,34}.

$$p^1(\tau_n) = [p_0^1(\tau_n) p_1^1(\tau_n) p_2^1(\tau_n) \dots p_8^1(\tau_n) p_9^1(\tau_n) p_{10}^1(\tau_n)] \quad (16)$$

$$\tau_n = n\Delta t \quad (17)$$

$$p^1(\tau_0) = [p_0^1(\tau_0) p_1^1(\tau_0) p_2^1(\tau_0) \dots p_8^1(\tau_0) p_9^1(\tau_0) p_{10}^1(\tau_0)] = [100 \dots 000] \quad (18)$$

According to Markov stochastic cellular models, a transition probability matrix, P^1 , containing values of transition probability P_{ji}^1 , is defined by Eq. (19)³⁰. P^1 can be determined using different probability density functions of random variable. The breaking of a cellulose species within the molecular mass class M_0^1 can occur in the position k in the chain, where $k=1, 2 \dots K-1$ and K is its polymerization degree (PD), resulting in two hydrolysates, one of them having $PD=k$, the other one $PD=K-k$. If $k=1$, the breaking of the hydrolysate with $PD=K-1$ can take place in the position k , where $k=1, 2 \dots K-2$. Accordingly, the values of random variable are $1, 2 \dots K-i$, where $i \geq 1$.

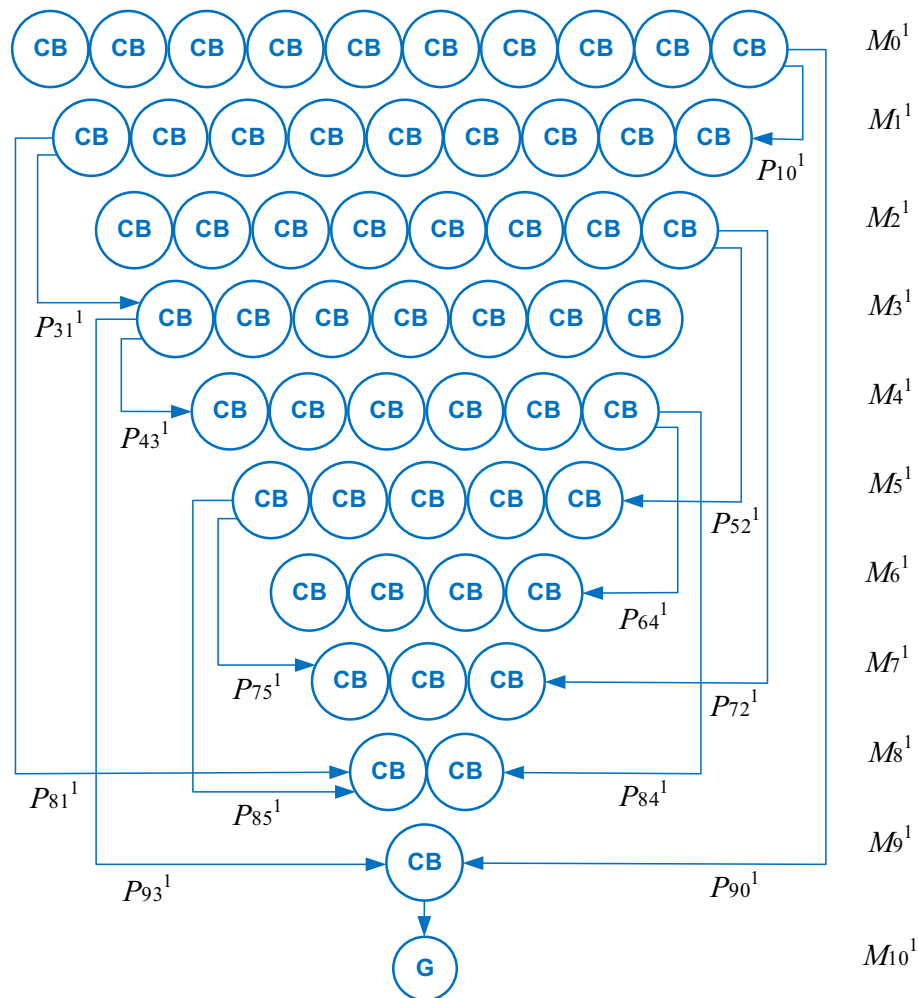


Figure 2. Hydrolysate molecular mass intervals and transition probabilities for a polysaccharide species (s=1) consisting of 10 units of cellobiose (CB).

$$P^1 = \begin{bmatrix} P_{00}^1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ P_{10}^1 & P_{11}^1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ P_{20}^1 & P_{21}^1 & P_{22}^1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ P_{30}^1 & P_{31}^1 & P_{32}^1 & P_{33}^1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ P_{40}^1 & P_{41}^1 & P_{42}^1 & P_{43}^1 & P_{44}^1 & 0 & 0 & 0 & 0 & 0 & 0 \\ P_{50}^1 & P_{51}^1 & P_{52}^1 & P_{53}^1 & P_{54}^1 & P_{55}^1 & 0 & 0 & 0 & 0 & 0 \\ P_{60}^1 & P_{61}^1 & P_{62}^1 & P_{63}^1 & P_{64}^1 & P_{65}^1 & P_{66}^1 & 0 & 0 & 0 & 0 \\ P_{70}^1 & P_{71}^1 & P_{72}^1 & P_{73}^1 & P_{74}^1 & P_{75}^1 & P_{76}^1 & P_{77}^1 & 0 & 0 & 0 \\ P_{80}^1 & P_{81}^1 & P_{82}^1 & P_{83}^1 & P_{84}^1 & P_{85}^1 & P_{86}^1 & P_{87}^1 & P_{88}^1 & 0 & 0 \\ P_{90}^1 & P_{91}^1 & P_{92}^1 & P_{93}^1 & P_{94}^1 & P_{95}^1 & P_{96}^1 & P_{97}^1 & P_{98}^1 & P_{99}^1 & 0 \\ P_{100}^1 & P_{101}^1 & P_{102}^1 & P_{103}^1 & P_{104}^1 & P_{105}^1 & P_{106}^1 & P_{107}^1 & P_{108}^1 & P_{109}^1 & 1 \end{bmatrix} \tag{19}$$

Assuming the highest breaking probability in the centre of polymeric chain, a Gauss (normal) distribution can be used to compute the probabilities in the transition probability matrix. A random variable X that takes values x is normally (Gauss) distributed when the probability density function is expressed by Eq. (20), where μ is the mean of x values, σ^2 the variance, and σ the standard deviation.

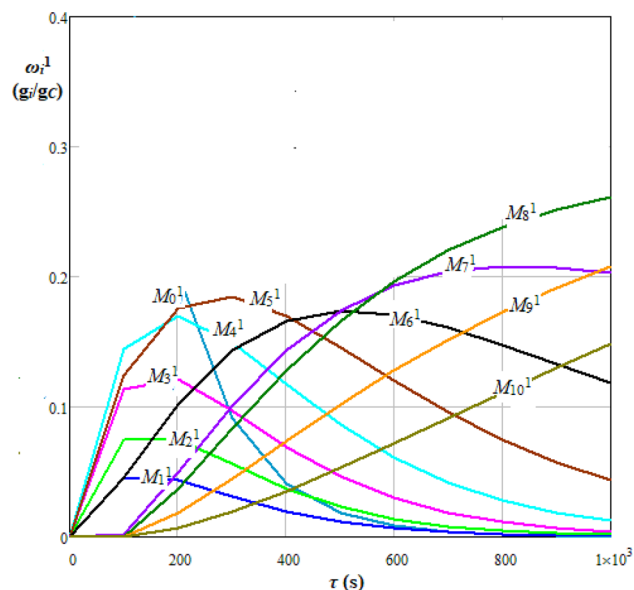


Figure 3. Dynamics of mass fraction of i hydrolysable species within the molecular mass class M_i^1 ($i=0, 1 \dots 10$), ω_i^1 , for Gauss (normal) distribution of transition probabilities.

$$f_{GS}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (20)$$

Normal distributions have many useful properties, so random variables with unknown distributions are often considered to be normal, especially in chemistry, physics, biology, etc. This assumption is based on the central limit theorem (CLT), which states that if sufficiently large random samples are taken from a population having any distribution of a variable and finite values of mean and variance, then the distribution of sample mean will approach a Gauss (normal) distribution. There is an obvious tendency in sciences and social life to assume normal distributions in applications where they may not be suitable. As Lippmann concluded, “Everybody believes in the exponential law of errors: the experimenters, because they think it can be proved by mathematics; and the mathematicians, because they believe it has been established by observation”³⁵.

Considering the mean (μ) and standard deviation (σ) given by Eqs. (21) and (22), the elements of transition probability matrix, P_{ji}^1 , based on a Gauss distribution can be determined by Eq. (23). For the hydrolysis case depicted in Fig. 2 ($i=0, 1 \dots N, N=10$), P^1 is given by Eq. (24). Dynamics of mass fraction of i hydrolysable species within the molecular mass class M_i^1 ($i=0, 1 \dots 10$), $\omega_i^1(\tau)$, which were obtained based on Eqs. (13)–(15) and (24), where $\Delta\tau=100$ s, are shown in Fig. 3.

$$\mu = \mu(x) = \frac{(N+1-x)}{2} \quad (21)$$

$$\sigma = \sigma(x) = \sqrt{\frac{(N+1-x)^2 - 1}{12}} \quad (22)$$

$$P_{ji}^1 = \begin{cases} \frac{1}{N+1} & \text{if } i=j=0 \\ 1 & \text{if } i=j=N \\ PP_{ji}^1 = \begin{cases} f_{GS}(j-i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(j-i-\mu)^2}{2\sigma^2}\right) & \text{if } i < j \\ 0 & \text{otherwise} \end{cases} & \text{if } i \neq j \\ 1 - \sum_{j=1}^N PP_{ji}^1 & \text{otherwise} \end{cases} \quad (23)$$

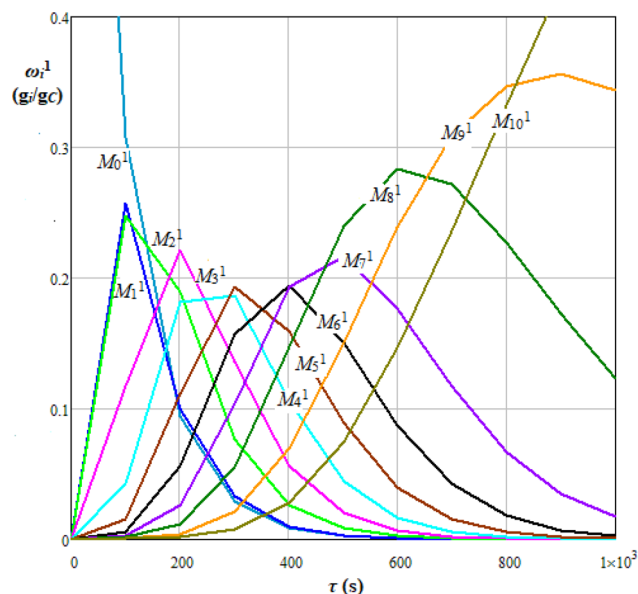


Figure 4. Dynamics of mass fraction of i hydrolysable species within the molecular mass class M_i^1 ($i=0, 1 \dots 10$), ω_i^1 , for Weibull distribution of transition probabilities.

$$P^1 = \begin{bmatrix} 0.318 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.036 & 0.352 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.066 & 0.036 & 0.352 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.114 & 0.066 & 0.036 & 0.352 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.170 & 0.114 & 0.066 & 0.036 & 0.357 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.176 & 0.170 & 0.114 & 0.066 & 0.036 & 0.437 & 0 & 0 & 0 & 0 & 0 \\ 0.080 & 0.176 & 0.170 & 0.114 & 0.066 & 0.036 & 0.613 & 0 & 0 & 0 & 0 \\ 0.005 & 0.080 & 0.176 & 0.170 & 0.114 & 0.066 & 0.036 & 0.783 & 0 & 0 & 0 \\ 0.000 & 0.005 & 0.080 & 0.176 & 0.170 & 0.114 & 0.066 & 0.036 & 0.898 & 0 & 0 \\ 0.000 & 0.000 & 0.005 & 0.080 & 0.176 & 0.170 & 0.114 & 0.066 & 0.036 & 0.964 & 0 \\ 0.000 & 0.000 & 0.000 & 0.005 & 0.080 & 0.176 & 0.170 & 0.114 & 0.066 & 0.036 & 1 \end{bmatrix} \quad (24)$$

Many dynamic phenomena in the chemical engineering field have been characterized and modelled using Weibull distribution^{36–38}. Probability density function of a Weibull random variable is given by Eq. (25), where $\beta > 0$ is the shape parameter, $\lambda > 0$ is the scale parameter, and $x \geq 0$ the values of Weibull random variable³⁹.

$$f_W(x) = \frac{\beta}{\lambda} \left(\frac{x}{\lambda}\right)^{\beta-1} \exp\left(-\left(\frac{x}{\lambda}\right)^\beta\right) \quad (25)$$

In the case of cellulose hydrolysis, the values of shape parameter are interpreted as follows: (i) $0 < \beta < 1$ indicates that the chain breaking rate decreases over time, meaning that the process hydrolysates become more stable, (ii) $\beta = 1$ implies a constant breaking rate over time, and (iii) $\beta > 1$ corresponds to an increase in the breaking rate. Characteristic parameters of Weibull distribution were expressed by Eqs. (26) and (27) depending on the integer values of Weibull random variable ($x \geq 0$), i.e., scale parameter (λ) was linked to the mean (μ) of random variable values and shape parameter ($\beta > 0.5$) to the variance (σ^2). For the hydrolysis case presented in Fig. 2, the elements of transition probability matrix, P_{ji}^1 , are given by Eq. (28). Dynamics of mass fraction of i hydrolysable species within the molecular mass class M_i^1 ($i=0, 1 \dots 10$), $\omega_i^1(\tau)$, obtained based on Eqs. (13)–(15), (26)–(28), where $\Delta\tau = 100$ s, are shown in Fig. 4.

$$\lambda = \lambda(x) = 1.4 + \frac{\mu(x)}{N} \quad (26)$$

$$\beta = \beta(x) = 0.95 + \frac{\sigma^2(x)}{N} \quad (27)$$

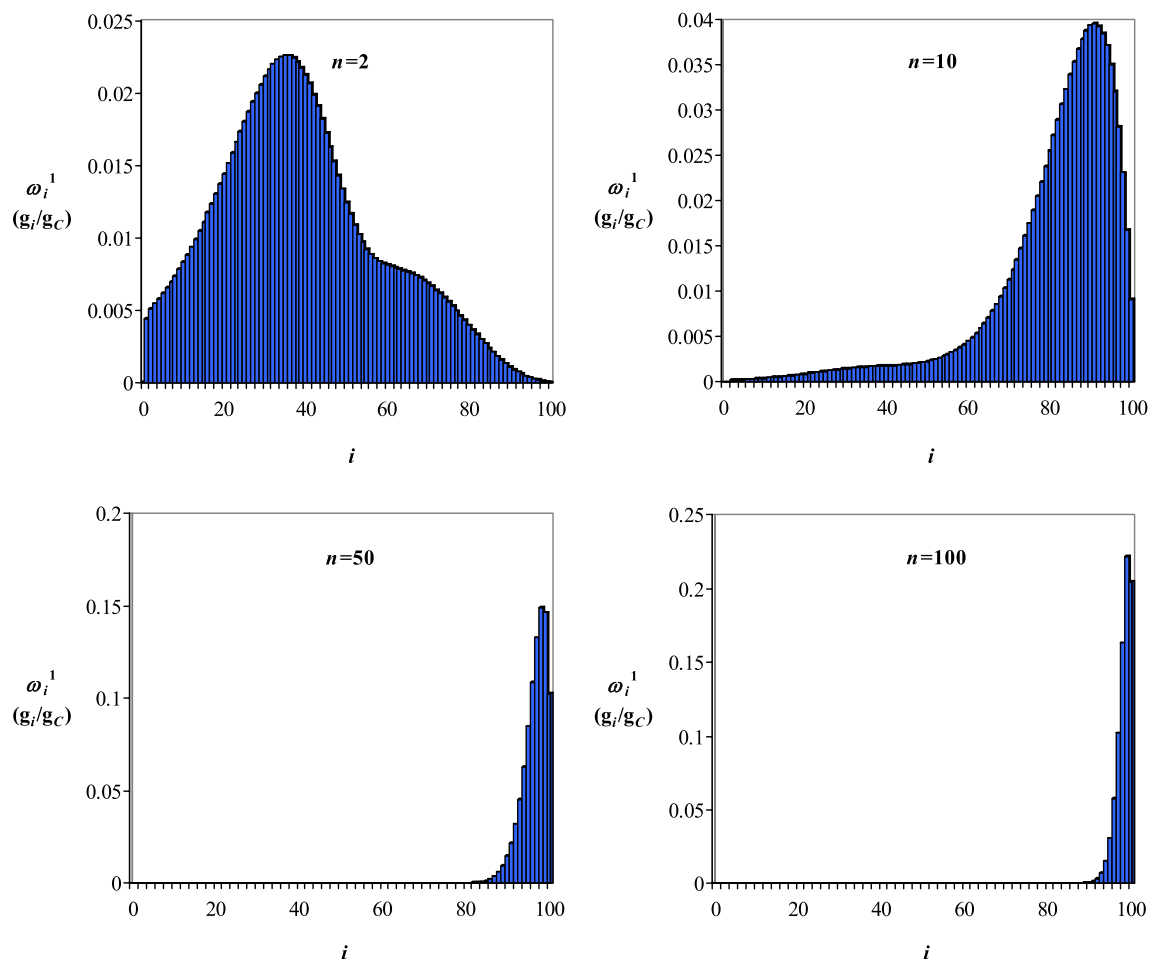
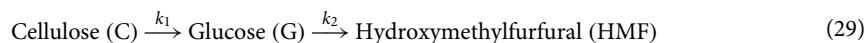


Figure 5. Mass fractions of i hydrolysable species within the molecular mass class M_i^1 ($i=0, 1 \dots 100$), ω_i^1 , at different time sequences ($n=2, 10, 50, 100$) for hydrolysis of a cellulose with $PD=200$ (100 units of CB) assuming a Gauss distribution of transition probabilities.

$$P_{ji}^1 = \begin{cases} \frac{3.5}{N+1} & \text{if } i = j = 0 \\ 1 & \text{if } i = j = N \\ PP_{ji}^1 = \begin{cases} f_W(j-i) = \frac{\beta}{\lambda} \left(\frac{j-i}{\lambda}\right)^{\beta-1} \exp\left(-\left(\frac{j-i}{\lambda}\right)^\beta\right) & \text{if } i < j \\ 0 & \text{otherwise} \end{cases} & \text{if } i \neq j \\ 1 - \sum_{j=1}^N PP_{ji}^1 & \text{otherwise} \end{cases} \quad (28)$$

Acid hydrolysis of cellulose consisting of 100 units of cellobiose. Mass fractions of i hydrolysable species, ω_i^1 (g_i/g_C), at different time sequences for acid hydrolysis of cellulose with a value of PD of 200 (corresponding to 100 units of CB), are shown in Figs. 5 and 6, where Gauss (Fig. 5) and Weibull (Fig. 6) distributions of transition probabilities were considered. The results presented in Figs. 5 and 6 highlight a relatively quick disappearance of macromolecules with high molecular mass, suggesting that the hydrolysis of species with low molecular mass is the rate-limiting step. This finding is in line with the conclusions of other researches, where the cellulose hydrolysis was assumed as a homogeneous kinetic process, occurred according to scheme (29) and described by Eqs. (30)–(32), where c_C , c_G , and c_{HMF} are the concentrations of cellulose (C), glucose (G), and hydroxymethylfurfural (HMF), k_1 and k_2 represent the rate constants, and τ is the time^{40–43}.



$$\frac{dc_C}{d\tau} = -k_1 c_C \quad (30)$$

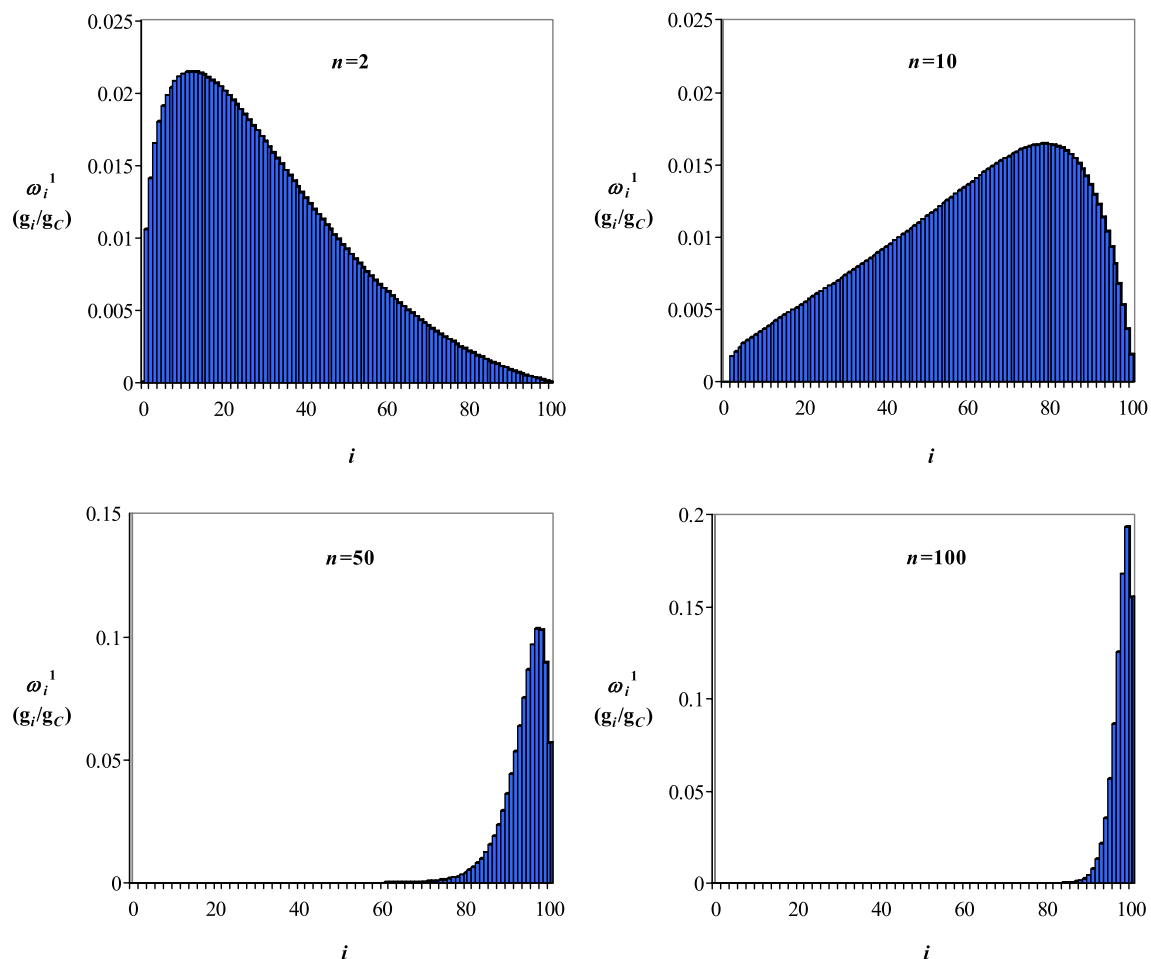


Figure 6. Mass fractions of i hydrolysable species within the molecular mass class M_i^1 ($i=0, 1 \dots 100$), ω_i^1 , at different time sequences ($n=2, 10, 50, 100$) for hydrolysis of a cellulose with $PD=200$ (100 units of CB) assuming a Weibull distribution of transition probabilities.

$$\frac{dc_G}{d\tau} = k_1 c_C - k_2 c_G \quad (31)$$

$$\frac{dc_{HMF}}{d\tau} = k_2 c_G \quad (32)$$

Acid hydrolysis dynamics are heavily affected by process temperature, acid type and concentration. Arrhenius equations [Eqs. (33) and (34)], where the contribution of acid concentration (divided by a reference value, $c_a/c_{a,ref}$) was taken into account, were used to express characteristic reaction rate constants of homogeneous kinetic model (k_1 and k_2) depending on absolute temperature (T).

$$k_1 = k_{10} \left(\frac{c_a}{c_{a,ref}} \right)^{n_1} \exp \left(-\frac{E_{A1}}{RT} \right) \quad (33)$$

$$k_2 = k_{20} \left(\frac{c_a}{c_{a,ref}} \right)^{n_2} \exp \left(-\frac{E_{A2}}{RT} \right) \quad (34)$$

According to Eqs. (33) and (34), transition probability frequencies in the stochastic model, α_{ji}^1 , can be estimated by Eq. (35), where E_{Am} represents a mean reaction activation energy (taking into account all decompositions to glucose), E_{AG} is the activation energy for transition from glucose (G) to HMF, and T_{ref} is a reference absolute temperature.

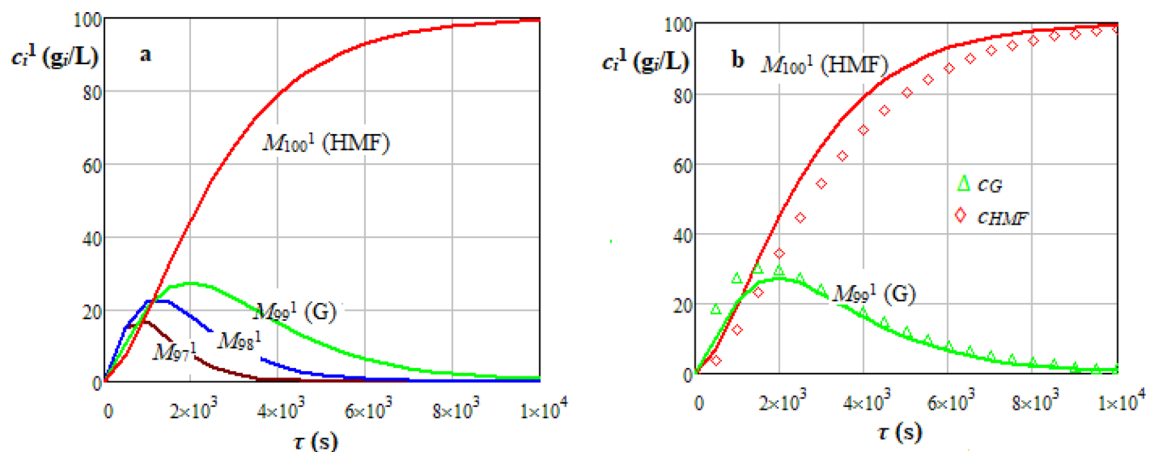


Figure 7. Dynamics of i hydrolysable species ($M_{97}^1 \dots M_{100}^1$) concentration, c_i^1 (g/L), predicted by stochastic model for sulphuric acid hydrolysis of a cellulose (C) suspension ($PD=200$, $c_{CO}=100$ g_C/L, $\Delta\tau=10$ s, Gauss distributed transition probabilities) (a) and comparison between stochastic model (line) and homogeneous deterministic kinetic model (points) for M_{99}^1 and M_{100}^1 ($T=483$ K, $T_{ref}=453$ K, $c_a=4\%$, $c_{a,ref}=1\%$) (b).

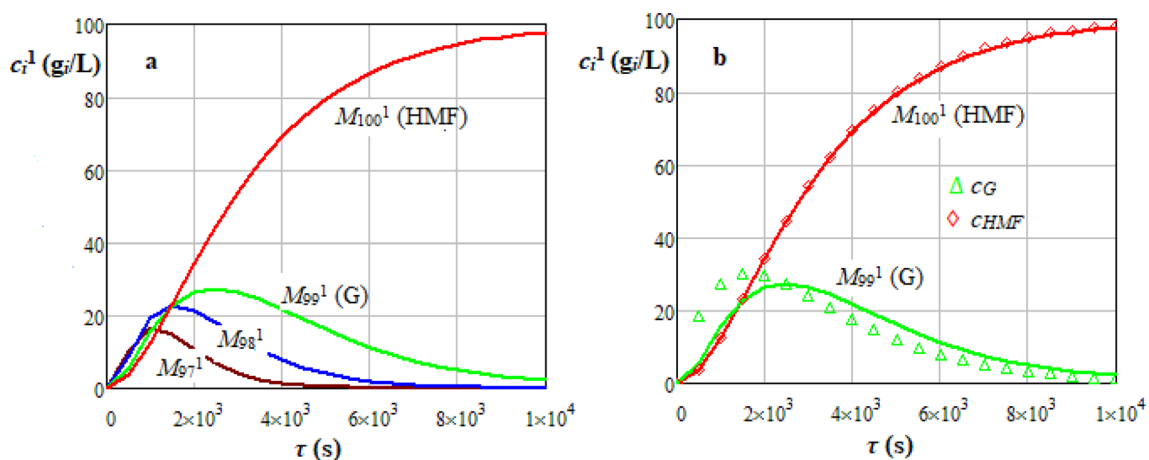


Figure 8. Dynamics of i hydrolysable species ($M_{97}^1 \dots M_{100}^1$) concentration, c_i^1 (g/L), predicted by stochastic model for sulphuric acid hydrolysis of a cellulose (C) suspension ($PD=200$, $c_{CO}=100$ g_C/L, $\Delta\tau=10$ s, Weibull distributed transition probabilities) (a) and comparison between stochastic model (line) and homogeneous deterministic kinetic model (points) for M_{99}^1 and M_{100}^1 ($T=483$ K, $T_{ref}=453$ K, $c=4\%$, $c_{a,ref}=1\%$) (b).

$$\alpha_{ji}^1 = \begin{cases} \alpha_{ji0}^1 \left(\frac{c_a}{c_{a,ref}} \right)^n \exp \left[\frac{E_{Am}}{R} \left(\frac{1}{T} - \frac{1}{T_{ref}} \right) \right], & i, j = 0, 1 \dots N - 1 \\ \alpha_{ji0}^1 \left(\frac{c_a}{c_{a,ref}} \right)^n \exp \left[\frac{E_{AG}}{R} \left(\frac{1}{T} - \frac{1}{T_{ref}} \right) \right], & i = N - 1, j = N \end{cases} \quad (35)$$

Dynamics of i hydrolysable species concentration, c_i^1 (g/L), predicted by stochastic model for cellulose hydrolysis are shown in Figs. 7a and 8a, where the matrix of transition probabilities, based on Gauss or Weibull distributions, was determined according to Eqs. (36)–(39). Dynamics of c_G and c_{HMF} predicted by a homogeneous deterministic kinetic model [Eqs. (30)–(34)] based on data reported in the related literature^{40–43} are represented in Figs. 7b and 8b. Values of characteristic parameters of deterministic and stochastic models are specified in Table S1⁵. The largest values of root mean square error (RMSE) and coefficient of variation (CV) for results presented in Figs. 7b and 8b are summarized in Table S2. Data predicted by deterministic and stochastic models for a cellulose with $PD=200$ assuming Gauss and Weibull distributed transition probabilities were in a reasonable agreement, i.e., $RMSE \leq 6.50$ g/L ($CV \leq 0.098$) and $RMSE \leq 5.26$ g/L ($CV \leq 0.410$).

$$P_{ji}^{-1} = \begin{cases} \frac{1}{N+1} & \text{if } i = j = 0 \\ 1 & \text{if } i = j = N \\ PP_{ji}^{-1} = \begin{cases} f_{GS}(j-i)g(c_a)h(T) & \text{if } i < j \\ 0 & \text{otherwise} \end{cases} & \text{if } i \neq j \\ 1 - \sum_{j=1}^N PP_{ji}^{-1} & \text{otherwise} \end{cases} \quad (36)$$

$$P_{ji}^{-1} = \begin{cases} \frac{3.5}{N+1} & \text{if } i = j = 0 \\ 1 & \text{if } i = j = N \\ PP_{ji}^{-1} = \begin{cases} f_W(j-i)g(c_a)h(T) & \text{if } i < j \\ 0 & \text{otherwise} \end{cases} & \text{if } i \neq j \\ 1 - \sum_{j=1}^N PP_{ji}^{-1} & \text{otherwise} \end{cases} \quad (37)$$

$$g(c_a) = \left(\frac{c_a}{c_{a,ref}} \right)^n \quad (38)$$

$$h(T) = \begin{cases} \exp \left[\frac{E_{Am}}{R} \left(\frac{1}{T} - \frac{1}{T_{ref}} \right) \right], & i, j = 0, 1, \dots, N-1 \\ \exp \left[\frac{E_{AG}}{R} \left(\frac{1}{T} - \frac{1}{T_{ref}} \right) \right], & i = N-1, j = N \end{cases} \quad (39)$$

Conclusions

The production of cellulosic biofuel, which is technologically feasible, expects the transition to industrial scale. Accordingly, the hydrolysis of cellulose and (ligno)cellulosic biomass has been of great interest lately. Numerous experimental and theoretical studies on acid hydrolysis have been reported in the related literature. Deterministic kinetic models are commonly applied to describe the acid hydrolysis.

Two stochastic models, one involving a Gauss distribution of transition probabilities, the other a Weibull distribution, were developed in this paper to predict the kinetics of acid hydrolysis of cellulose. It was assumed that the cellulose subjected to the hydrolysis contained a finite number of hydrolysable polysaccharide species (cellulose fragments with different PD), random breaking of a polysaccharide was based on the cleavage of a parent macromolecule into two descendants, and the molecular fragmentation was a homogeneous Markov process. Model equations, transition probability matrix and dynamics of mass fraction of hydrolysable species for a cellulose consisting of 10 units of CB ($PD=20$) were presented in Methods and the first part of Results and discussions. The models were then extended, considering the effects of acid concentration and process temperature on the kinetics of acid hydrolysis of a cellulose consisting of 100 units of CB ($PD=200$). Dynamics of mass concentration of hydrolysable species predicted by stochastic models and those determined using a homogeneous deterministic kinetic model were compared and a reasonable agreement was obtained. Both stochastic models can accurately predict the kinetics of acid hydrolysis and cover the limitations of some deterministic kinetic models, e.g., large number of equations and parameters, modification of parameter values by changing the process conditions.

Received: 22 December 2020; Accepted: 12 April 2021

Published online: 04 May 2021

References

- Bansal, P., Hall, M., Realff, M. J., Lee, J. H. & Bommarius, A. S. Multivariate statistical analysis of X-ray data from cellulose: A new method to determine degree of crystallinity and predict hydrolysis rates. *Bioresour. Technol.* **101**(12), 4461–4471. <https://doi.org/10.1016/j.biortech.2010.01.068> (2010).
- Brandam, C., Meyer, X. M., Proth, J., Strehaiano, P. & Pingaud, H. An original kinetic model for the enzymatic hydrolysis of starch during mashing. *Biochem. Eng. J.* **13**(1), 43–52. [https://doi.org/10.1016/S1369-703X\(02\)00100-6](https://doi.org/10.1016/S1369-703X(02)00100-6) (2003).
- Chen, C. C., Dai, L., Ma, L. & Guo, R. T. Enzymatic degradation of plant biomass and synthetic polymers. *Nat. Rev. Chem.* **4**(3), 114–126. <https://doi.org/10.1038/s41570-020-0163-6> (2020).
- Cho, E. J., Trinh, L. T. P., Song, Y., Lee, Y. G. & Bae, H. J. Bioconversion of biomass waste into high value chemicals. *Bioresour. Technol.* **298**, 122386. <https://doi.org/10.1016/j.biortech.2019.122386> (2019).
- Dobre, T., Părvulescu, O. C., Răducanu, C., Trică, B. & Jinescu, G. Stochastic modelling of polysaccharide hydrolysis. *J. Eng. Sci. Innov.* **3**(1), 25–38 (2018).
- Fenila, F. & Shastri, Y. Stochastic optimization of enzymatic hydrolysis of lignocellulosic biomass. *Comput. Chem. Eng.* **135**, 106776. <https://doi.org/10.1016/j.compchemeng.2020.106776> (2020).
- Jacobsen, S. E. & Wyman, C. E. Cellulose and hemicellulose hydrolysis models for application to current and novel pretreatment processes. *Appl. Biochem. Biotechnol.* **84**, 81–96. <https://doi.org/10.1385/ABAB:84-86:1-9:81> (2000).

8. Qiao, Y. *et al.* Thermal decomposition of castor oil, corn starch, soy protein, lignin, xylan, and cellulose during fast pyrolysis. *Bioresour. Technol.* **278**, 287–295. <https://doi.org/10.1016/j.biortech.2019.01.102> (2019).
9. Sarkar, A. K. & Eters, J. N. Enzymatic hydrolysis of cotton fibers: Modelling using empirical equations. *Journal of Cotton Science* **8**(4), 254–260 (2004).
10. Torget, R. W., Kim, J. S. & Lee, Y. Y. Fundamental aspects of dilute acid hydrolysis/fractionation kinetics of hardwood carbohydrates. 1. Cellulose hydrolysis. *Ind. Eng. Chem. Res.* **39**(8), 2817–2825. <https://doi.org/10.1021/ie990915q> (2000).
11. Girisuta, B., Janssen, L. P. B. M. & Heeres, H. J. Kinetic study on the acid-catalyzed hydrolysis of cellulose to levulinic acid. *Ind. Eng. Chem. Res.* **46**(6), 1696–1708. <https://doi.org/10.1021/ie061186z> (2007).
12. Kadam, K. L., Rydholm, E. C. & McMillan, J. D. Development and validation of a kinetic model for enzymatic saccharification of lignocellulosic biomass. *Biotechnol. Prog.* **20**(3), 698–705. <https://doi.org/10.1021/bp034316x> (2004).
13. Nidetzky, B. & Steiner, W. A new approach for modeling cellulase-cellulose adsorption and the kinetics of the enzymatic hydrolysis of microcrystalline cellulose. *Biotechnol. Bioeng.* **42**(4), 469–479. <https://doi.org/10.1002/bit.260420410> (1993).
14. Saeman, J. F. Kinetics of wood saccharification: Hydrolysis of cellulose and decomposition of sugars in dilute acid at high temperature. *Ind. Eng. Chem.* **37**(1), 43–52. <https://doi.org/10.1021/ie50421a009> (1945).
15. Xiang, Q., Lee, Y. Y., Pettersson, P. O. & Torget, R. W. Heterogeneous aspects of acid hydrolysis of α -cellulose. *Appl. Biochem. Biotechnol.* **107**(1–3), 505–514 (2003).
16. Bansal, P., Hall, M., Realf, M. J., Lee, J. H. & Bommarius, A. S. Modeling cellulase kinetics on lignocellulosic substrates. *Biotechnol. Adv.* **27**(6), 833–848. <https://doi.org/10.1016/j.biotechadv.2009.06.005> (2009).
17. Turon, X., Rojas, O. J. & Deinhammer, R. S. Enzymatic kinetics of cellulose hydrolysis: A QCM-D study. *Langmuir* **24**(8), 3880–3887. <https://doi.org/10.1021/la7032753> (2008).
18. Zhang, Y. H. P. & Lynd, L. R. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems. *Biotechnol. Bioeng.* **88**(7), 797–824. <https://doi.org/10.1002/bit.20282> (2004).
19. Drissen, R. E. T. *et al.* A generic model for glucose production from various cellulose sources by a commercial cellulase complex. *Biocatal. Biotransform.* **25**(6), 419–429. <https://doi.org/10.1080/10242420701510668> (2007).
20. Liang, C., Gu, C., Raftery, J., Karim, N. M. & Holtzapfle, M. Development of modified HCH-1 kinetic model for long-term enzymatic cellulose hydrolysis and comparison with literature models. *Biotechnol. Biofuels* **12**, 34. <https://doi.org/10.1186/s13068-019-1371-5> (2019).
21. Sousa, R. Jr., Carvalho, M. L., Giordano, R. L. C. & Giordano, R. C. Recent trends in the modeling of cellulose hydrolysis. *Braz. J. Chem. Eng.* **28**(4), 545–564. <https://doi.org/10.1590/S0104-66322011000400001> (2011).
22. Wojcik, M. & Milek, J. A new method to determine optimum temperature and activation energies for enzymatic reactions. *Bioprocess Biosyst. Eng.* **39**, 1319–1323. <https://doi.org/10.1007/s00449-016-1596-7> (2016).
23. Yang, B., Dai, Z., Ding, S. Y. & Wyman, C. E. Enzymatic hydrolysis of cellulosic biomass. *Biofuels* **2**(4), 421–449. <https://doi.org/10.4155/bfs.11.116> (2011).
24. Marongiu, A., Faravelli, T., Bozzano, G., Dente, M. & Ranzi, E. Thermal degradation of poly(vinyl chloride). *J. Anal. Appl. Pyrolysis* **70**(2), 519–553. [https://doi.org/10.1016/s0165-2370\(03\)00024-x](https://doi.org/10.1016/s0165-2370(03)00024-x) (2003).
25. Mehl, M. *et al.* A kinetic modeling study of the thermal degradation of halogenated polymers. *J. Anal. Appl. Pyrolysis* **72**(2), 253–272. <https://doi.org/10.1016/j.jaap.2004.07.007> (2004).
26. Staggs, J. E. J. Modelling random scission of linear polymers. *Polym. Degrad. Stab.* **76**(1), 37–44. [https://doi.org/10.1016/s0141-3910\(01\)00263-4](https://doi.org/10.1016/s0141-3910(01)00263-4) (2002).
27. Kostoglou, M. Mathematical analysis of polymer degradation with chain-end scission. *Chem. Eng. Sci.* **55**(13), 2507–2513. [https://doi.org/10.1016/s0009-2509\(99\)00471-6](https://doi.org/10.1016/s0009-2509(99)00471-6) (2000).
28. Kumar, D. & Murthy, G. S. Stochastic molecular model of enzymatic hydrolysis of cellulose for ethanol production. *Biotechnol. Biofuels* **6**, 63. <https://doi.org/10.1186/1754-6834-6-63> (2013).
29. Nassar, R., Chou, S. T. & Fan, L. T. Stochastic analysis of stepwise cellulose degradation. *Chem. Eng. Sci.* **46**(7), 1651–1657. [https://doi.org/10.1016/0009-2509\(91\)87012-2](https://doi.org/10.1016/0009-2509(91)87012-2) (1991).
30. Dobre, T. & Sanchez Marcano, J. Stochastic modelling. Chapter 4 in *Chemical Engineering: Modelling, Simulation, and Similitude*. 191–322 (Wiley VCH, 2007).
31. Dymarski, P. (Ed.) *Hidden Markov Models, Theory and Applications*. InTechOpen. <http://www.intechopen.com> (2011).
32. Gaspard, P. & Andrieux, D. Kinetics and thermodynamics of first-order Markov chain copolymerization. *J. Chem. Phys.* **141**, 044908. <https://doi.org/10.1063/1.4890821> (2014).
33. Gillespie, D. T. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* **58**(1), 35–55. <https://doi.org/10.1146/annurev.physchem.58.032806.104637> (2007).
34. Varrat, A. *et al.* Structural basis for ligand binding and processivity in cellobiohydrolase Cel6A from *Humicola insolens*. *Structure* **11**(7), 855–864. [https://doi.org/10.1016/S0969-2126\(03\)00124-2](https://doi.org/10.1016/S0969-2126(03)00124-2) (2003).
35. Whittaker, E. T. & Robinson, G. *Normal Frequency Distribution*. Chapter 8 in *The Calculus of Observations: A Treatise on Numerical Mathematics* 164–208 (Dover, 1967).
36. Khorshadizadeh, M. & Atashi, H. Modeling the kinetics of cobalt Fischer-Tropsch catalyst deactivation trends through an innovative modified Weibull distribution. *Phys. Chem. Chem. Phys.* **29**, 19252–19261. <https://doi.org/10.1039/C7CP02210G> (2017).
37. Kolar-Anić, L., Veljković, S., Kapor, S. & Dubljević, B. Weibull distribution and kinetics of heterogeneous processes. *J. Chem. Phys.* **63**(2), 663–668. <https://doi.org/10.1063/1.431388> (1975).
38. Mafart, P., Couvert, O., Gaillard, S. & Leguerinel, I. On calculating sterility in thermal preservation methods: Application of the Weibull frequency distribution model. *Int. J. Food Microbiol.* **72**(1–2), 107–113. [https://doi.org/10.1016/s0168-1605\(01\)00624-9](https://doi.org/10.1016/s0168-1605(01)00624-9) (2002).
39. Murthy, D. N. P., Xie, M. & Jiang, R. *Weibull Models* (Wiley, 2003).
40. Lenihan, P. *et al.* Dilute acid hydrolysis of lignocellulosic biomass. *Chem. Eng. J.* **156**(2), 395–403. <https://doi.org/10.1016/j.cej.2009.10.061> (2010).
41. Morales-delaRosa, S., Campos-Martin, J. M. & Fierro, J. L. G. Optimization of the process of chemical hydrolysis of cellulose to glucose. *Cellulose* **21**, 2397–2407. <https://doi.org/10.1007/s10570-014-0280-9> (2014).
42. Orozco, A., Ahmad, M., Rooney, D. & Walker, G. Dilute acid hydrolysis of cellulose and cellulosic bio-waste using a microwave reactor system. *Process Saf. Environ. Prot.* **85**(5), 446–449. <https://doi.org/10.1205/psep07003> (2007).
43. Xiang, Q., Lee, Y. Y. & Torget, R. W. Kinetics of glucose decomposition during dilute-acid hydrolysis of lignocellulosic biomass. In *Proceedings of the Twenty-Fifth Symposium on Biotechnology for Fuels and Chemicals* (eds Finkelstein, M. *et al.*) (Humana Press, 2004). https://doi.org/10.1007/978-1-59259-837-3_91.

Author contributions

J.M.D., O.C.P. and T.D. wrote the main manuscript text; O.C.P. had a substantial contribution to the conception of article and interpretation of results; T.D. had a substantial contribution to the stochastic models; C.E.R. prepared Figs. 1 and 2. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-88873-6>.

Correspondence and requests for materials should be addressed to O.C.P.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021