Multiple-distribution DAEM modelling of spruce pyrolysis: An investigation of the best trade-off regarding the number and shape of distributions

Yong Tian, Patrick Perré

Multiple-distribution DAEM modelling of spruce pyrolysis: an investigation of the best trade-off regarding the number and shape of distributions

Yong TIAN[a,*], Patrick Perré[a,b]

[a] LGPM, CentraleSupélec, Université Paris-Saclay, 3 Rue Joliot Curie, 91190 Gif-sur-Yvette, France

[b] LGPM, Centre Européen de Biotechnologie et de Bioéconomie (CEBB), 3 Rue des Rouges Terres, 51110, Pomacle, France

[*]Corresponding author.

E-mail address: Yong.Tian@centralesupelec.fr (Yong TIAN), patrick.perre@centralesupelec.fr (Patrick Perré)

Abstract

The distributed activation energy model (DAEM) is a widely used, accurate and robust method to model biomass pyrolysis. However, the appropriate numerical strategy in terms of distribution number and shape has not been systematically determined. This study analysed spruce powder pyrolysis under different scenarios of multiple-distribution DAEMs with symmetric/asymmetric distributions (Gaussian, logistic and exponential) and different distribution numbers. Dynamic tests at four heating rates (1, 2, 5 and 10 °C/min up to 800 °C) provided solid numerical learning database, and the optimization of residues between numerical calculation and database enabled identification of model parameters. Subsequently, validation was performed with static tests (250 to 500 °C with an interval of 50 °C and 2h-isothermal stages), and their corresponding residue analysis provided a fundamental basis to assess the model's true prediction ability. The trade-off between the model's prediction ability and degrees of freedom was robustly investigated with regard to the number and shape of the distribution. As stated by the quality of validations, a series of Gaussian-DAEMs (distribution number ranged from one to five) allowed for the determination of the best trade-off when the distribution number was three. Finally, the two-Gaussian plus one exponential distribution

exhibited the best overall prediction capacity among different multiple-distribution DAEMs, and was confirmed as the best strategy with regard to both distribution shape and number. A DTG simulation investigated each model's simulation effects with three assigned variation sections and justified the correspondence between pseudo-components and biomass constituents. Finally, the DAEM's capability to distinguish the effects of heating rate was demonstrated.

Keywords: Pyrolysis, DAEM, Exponential distribution, Gaussian distribution, Trade-off

1. Introduction

Biomass pyrolysis receives substantial attention as a key thermal conversion technology in bio-refineries [1] in the production of high value-added chemicals (bio-oil, char and biogas). It also acts as the main initial process for gasification and combustion. The ability to predict constituent kinetics and thermal behaviours during pyrolysis are therefore of great importance, namely for reactor design and industrial scale-up. Thermogravimetric analysis (TGA) is an advanced technology used to study the pyrolysis mechanism relating to devolatilization, which also provides sufficient data for kinetic modelling [2].

Pyrolysis is a rather complicated process with numerous chemical reactions, traditional methods of global first- and second-order kinetics are not applicable to biomass pyrolysis [3], which encourages the development of more effective models. Among them, a lumped-kinetic model accounting for more than 30 species has been developed [4], and competitive multi-step models have helped include primary devolatilizations and homogeneous secondary reactions in biomass components [5]. The distributed activation energy model (DAEM) has been proposed to manage the complexity of chemical reactions [6]. In this approach, pyrolysis is assumed to proceed as independently parallel reactions with different activation energies, which are further described by distribution functions.

The DAEM has been proposed to be the most comprehensive model to represent the pyrolysis of various complex feedstocks including coal [7], oil shale [8], sewage sludge [9], and biomass [6]. Its prediction kinetics are believed capable of working as sub-models in further industrial simulations [10]. Multiple-distribution DAEM is able to represent major chemical constituents [11] or multiple reaction stages [5], contributing to precise simulations of kinetics. Three-distribution models are widely applied due to its comprehensiveness and excellent correspondence with hemicellulose, cellulose and lignin [12]. Meanwhile the choices of two [13], four [14] and five [15] distributions are also proposed.

Concerning the mathematical forms of distribution, continuous statistic distributions are normally applied [15]. Meanwhile finite discrete distribution is also employed, which introduces relations between activation energy and pre-exponential factor either with [16] or without functional forms [17]. Yet differences between these two distribution types could be basically limited, since continuous distributions are necessarily discretized for computational implementations.

While symmetric distributions such as Gaussian and logistic have been widely applied in DAEMs [18], asymmetric distributions have also attracted attentions, as partial reactivity distribution in pyrolysis tends to be asymmetric, notably during the final stages [6]. Asymmetric distributions, including Weibull and gamma, have been assessed in several studies. Lakshmanan [19] first employed a Weibull DAEM to describe the thermal-chemical kinetics of multiple types of biomass. Recently, Li [20] found the Weibull outperformed other distributions in a study of two- and three-distribution DAEMs in polymer pyrolysis testing of both symmetric and asymmetric distributions. Xu [21] compared single-distribution DAEMs to asymmetric gamma, Rayleigh and Weibull distributions, and found that kinetic parameters relied heavily on distribution form. Alok [22] used asymptotic expansion for gamma distribution's numerical integral in a DAEM, yet the simulation effect was poor. Gamma

distribution offers a wide range of shapes that are capable to fit various kinetic profiles. In the decomposition context, its rate parameter measures the average life-time of active component [23], while its shape parameter endows multiple forms. Exponential distribution, as the degeneracy of gamma, shows particular features in lifetime distribution, stochastic process in general [24] and reliability analysis [25]. However, to the best of our knowledge, it has never been employed in multiple-distribution DAEM.

The application of multiple-distribution DAEMs has usually focused on the corresponding distribution number with equal pseudo-components or multiple-stage processes [26, 27], yet the effects of distribution number on prediction ability and numerical complexity have received little attention. Indeed, the increase of the distribution number in DAEMs could improve accuracy [15], as more subtle details might be captured in addition to the reaction kinetics. However, the simultaneous growth in numerical complexity could substantially endanger model robustness. Furthermore, the determination of parameters would face local minima or even be meaningless after model identification [28]. On the other hand, single-distribution DAEM shows insufficiency for biomass pyrolysis. In particular, single-Gaussian was found inappropriate to reproduce DTG data [15], and more than one logistic distribution was required for kinetic description [29]. Therefore, regarding the choice of distribution number, a trade-off between a model's prediction capacity and degrees of freedom should be seriously considered.

In previous studies, kinetic parameters have generally been determined by single non-isothermal experiments, which introduce the risks of local minima or compensation effects [18]. In this context, multiple experimental data have been proposed to reduce parameter uncertainties, especially those of activation energy and pre-exponential factors [30]. Pyrolysis profiles with two or more heating rates have been proved effective to distinguish between

kinetic models [3], and they could mitigate the compensation effect and more closely resemble operations in a genuine industrial system.

Generally, identification of model parameters has been based on the principle of minimizing residues between numerical calculations and the learning database, during which advanced optimization algorithms are applied such as the pattern search method [31], differential evolution algorithm [32] and genetic algorithm [33]. Beyond the pursuit of high accuracy and efficiency during model identification, the more important model validations should be emphasized to assess the model's true fit qualities under different reaction conditions. Várhegyi [34] evaluated prediction ability at 40 °C/min with model determined at 4 °C/min. Scott [35] extrapolated kinetic parameters that were identified at 20 °C/min and 30 °C/min to the theoretical curve at 10000 °C/min. Lin [36] performed predictions for 15 °C/min and 25 °C/min with the model that was identified at 20 °C/min. Nonetheless, solid validations with quantifiable uncertainties were relatively rare. Only recently, Ahmad [37] applied artificial neural network (ANN) for validating DAEM accuracy with histogram error distribution. And error analyses, in forms of absolute percentage error (MAPE) and root mean square error (RMSE), were employed in predicting mass loss at different heating rate [38] and biomass type [39] for validation purposes. On the other hand, the kinetic triplets (activation energy, pre-exponential factor and reaction rate) are barely verified since the multiple temperature profiles of parameter determinations couldn't be identical for prediction tests [40]. However, to assess the true model applicability in pyrolysis kinetics, it is logical and necessary to test the fitted parameters against reaction conditions that are different from those used for identification. Therefore, validation with additional temperature profiles should be conducted.

This work aims to assess the applicability of the multiple-distribution DAEM considering distribution shape (symmetry/asymmetry), and distribution number using a

rigorous approach. To that purpose, a set of dynamic tests were used as learning database and a completely different dataset, consisting of static tests over a wide range of plateau temperatures, was used as a validation database. The best trade-offs between the number of degrees of freedom and the prediction quality will be determined by using up to five-Gaussian distributions. The choice of distribution shape will be tested with two extra DAEMs (three-logistic and two Gaussian + one exponential). Performance of the distributions and its correspondence with biomass constituents will be studied in the subsequent DTG simulations. Finally, DAEM's performance in distinguishing the effect of the heating rate will be analysed.

2. Material and methods

2.1 Material

The biomass used in this study is European spruce (*Picea abies*), a softwood species. A 73-year-old tree was originally cut from the Auvergne region, France and subsequently processed to samples. A tree log 40–50 cm in diameter, 2 m in length was cut 2 m above the bottom. It was cut axially into 2.5-cm thick boards, and a portion 10 cm from the centre was taken to make samples for pyrolysis analysis. A rectangular column 2.5×2.5×5 cm^3 was cut from the healthy sapwood part of the board, where wood properties were relatively uniform. It was first sliced and ground in a cutting mill (RETSCH SM300) with a bottom sieve of 1 mm trapezoidal holes, followed by additional grinding with a universal mill (M20-IKA). A sieve stack of 0.063 mm and 0.08 mm opening sizes was used for sieving wood powder in a vibratory sieve shaker (RETSCH AS 200) at an amplitude of 90 % for 30 min. The sieved wood powder between 0.063–0.08 mm was dried at 105 °C for 24 h and stored in a desiccator.

Table 1 lists the basic chemical information of sprue sample. The ultimate and proximate analyses of the wood sample (density 450 kg/m^3) on a dry basis were conducted using a Thermo Fisher Scientific FLASH 2000 organic elementary analyser and Nabertherm

153 LV/9/11 furnace, respectively, following the ASTM E1755 and E872 standards. The chemical
154 composition analysis was based on the standard method of NREL (National Renewable
155 Energy Laboratory, U.S. Department of Energy).

156 Table 1. Results of proximate, ultimate and chemical composition analyses of European
157 spruce powder on a dry basis

Proximate analysis (wt.%)			Ultimate analysis (wt.%)			Chemical composition (wt.%)		
Volatile	Ash	FC	C	H	O	Cellulose	Hemicellulose	Lignin
84.93	0.29	14.78	47.1	6.0	43.7	42.49	14.89	31.59

158
159

160 2.2 Experimental method

161 Pyrolysis of spruce powder was performed in a thermogravimetric analyser (TGA, STA
162 449 F3 Jupiter, NETZSCH). TG signals were detected at data acquisition intervals of 0.1 min.
163 For each test, a ca. 10-mg sample was evenly spread in an alumina crucible. Measurements
164 were conducted under a pure nitrogen (99.999%) purge and protective gases at 50 ml/min and
165 20 ml/min, respectively.

166 Dynamic tests consisted of four different heating rates (1 °C/min, 2 °C/min, 5 °C/min
167 and 10 °C/min) during the pyrolysis stages. The entire temperature program started by
168 increasing the temperature from 30 °C to 100 °C at 10 °C/min, then maintaining it for 30 min
169 to eliminate the residual water presented in sample. The temperature was linearly increased to
170 800 °C at the four aforementioned heating rates, then cooled to room temperature under the
171 nitrogen purge. The thermogravimetric data from dynamic tests were set as a learning
172 database for the subsequent model identification process.

173 Static tests served as a model validation database. Similar to the dynamic tests above, the
174 wood sample was heated from 30 °C to 100 °C at 10 °C/min and held for 30 min. Then the
175 temperature was raised to the plateau at 10 °C/min, and an isothermal period maintained at

this value for 2 h. A series of tests were performed from 250 °C to 500 °C in 50 °C increments (250, 300, 350, 400, 450, 500 °C).

Temperature and sensitivity calibrations were performed in advance with standard materials specific to crucible type, temperature rate and gas type. The certified standards (NETZSCH calibration set) include indium, tin, bismuth, zinc, aluminium and silver. A blank was analysed before every test with the same crucible to exclude buoyancy effects and thermal drift. Dimensionless residual mass (DRM) and conversion rate (X_{exp}) were used to manage TG signals as:

$$DRM(t) = \frac{m_t}{m_0} \times 100\%, \qquad X_{exp}(t) = 1 - DRM(t) \tag{1}$$

in which m_t is the remaining mass at time t, and m_0 the dry mass, determined as the mass after the 30-minute plateau at 100 °C.

All dynamic and static tests were repeated twice to ensure accuracy. Standard deviations (SD) of all DRM data between two duplicate tests were calculated for verifications, herein table 2 lists two indexes: mean and maximum values of SD. The mean values got very limited range between 0.12% to 0.48%, and the maximum values that represented severe situation, was ranged between 0.30% to 1.81%. They both provided solid proofs of small errors between two duplicates and ensured the repeatability of experimental data.

Table 2. Mean and maximum values of standard deviations (SD) of DRM values in duplicate tests

Values of SD	Static test						Dynamic test			
	250 °C	300 °C	350 °C	400 °C	450 °C	500 °C	1 °C/min	2 °C/min	5 °C/min	10 °C/min
Mean (%)	0.46	0.13	0.12	0.47	0.48	0.29	0.33	0.26	0.41	0.32
Maximum (%)	0.88	0.30	1.61	1.12	0.73	0.44	1.30	0.98	1.32	1.81

3. Model formulation

3.1 DAEM formulation

The distributed activation energy model (DAEM) treats biomass pyrolysis as numerous parallel and irreversible first-order reactions, among them, decomposition rate of reaction i is：

$$\frac{dV_i(t)}{dt} = k_i\left(V_i^{\infty} - V_i(t)\right) \tag{2}$$

in which V_i^{∞} represents maximum volatile production from reaction i, and $V_i(t)$ is the generated volatile at time t. Reaction rate constant k_i is defined by the Arrhenius equation with the pre-exponential factor (A_i) and activation energy (E_i):

$$k_i = A_i exp\left(-\frac{E_i}{RT(t)}\right) \tag{3}$$

in which R is a universal gas constant and $T(t)$ is the temperature at time t. The compensation effects could provide different but equally good-fit sets of parameters, bringing much inaccuracy in identifications. Therefore, the value of pre-exponential factor is usually fixed to avoid ill-conditioned parameters, meantime being consistent with the transition-state theory (A≈10^11-10^16 s^(-1)) [34]. Here, A_i is assumed as constant (A) for all reactions. $V(t)$ is the total volatile production at time t, and statistical distribution $f(E)$ describes the activation energy, resulting in the integral form of conversion degree $X(t)$:

$$X(t) = \frac{V(t)}{V^{\infty}} = 1 - \int_0^{\infty} exp\left(-\int_{t_0}^{t} Ae^{-E/RT(t)}dt\right) f(E)dE \quad (4)$$

Wood is treated as the sum of multiple pseudo-components without any interactions during pyrolysis. Distributions $f_j(E)$ with a different weighting factor $V_{max}(j)$ are assigned to the pseudo-component j in wood ($j \in [1: N_d]$), N_d is the total number of distributions. Superposition of their volatile productions provides the final formulation of biomass conversion degree as:

$$X(t) = 1 - \sum_{j=1}^{N_d} V_{max}(j) \int_0^{\infty} exp\left(-A\int_0^{t} e^{-E/RT(t)}dt\right) f_j(E)dE \quad (5)$$

3.2 Distribution functions and mathematical implementation

Two common symmetric distributions, Gaussian and logistic distributions, and one asymmetric distribution, gamma distribution, were used in this study. These distributions in DAEM could represent physically the atomic interactions caused by the variability of macromolecules and their interaction in the cell wall [41]. Table 3 summarizes their mathematical formulas of probability density function (PDF), mean value and standard deviation. Their representative curves are depicted in Figure 1. Gaussian distribution, also known as normal distribution, is formulated by its mean value (E_0) and standard deviation value (σ). Its PDF curve is symmetrical and bell-shaped. Logistic distribution is another important symmetric probability distribution. It resembles Gaussian distribution but has heavier tails (larger kurtosis value) and thinner peaks around the mean value. The gamma distribution is defined by the shape parameter α and the rate parameter β. A minimum activation energy (E_{min}) must be added to obtain sufficient degrees of freedom to define the kinetics. The factor $\frac{E-E_{min}}{E_{min}}$ therefore scales the gamma distribution. The flexibility of the gamma distribution can produce different curve shapes by suitable combinations of its two parameters. For instance, when $\alpha = 1$, the gamma distribution becomes an exponential

distribution; when $\alpha = \frac{k}{2}$ (k is nature number) and $\beta = 0.5$, it becomes a Chi-squared distribution.

Table 3. Three types of distribution used in the research

Distributions	PDF	Mean	Standard deviation
Gaussian	$\mathrm{f(E)} = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{(E-E_0)^2}{2\sigma^2}\right]$	E_0	σ
Logistic	$\mathrm{f(E)} = \frac{\pi}{\sqrt{3}\sigma} \frac{exp\left[-\frac{\pi(E-E_0)}{\sqrt{3}\sigma}\right]}{\left\{1+exp\left[-\frac{\pi(E-E_0)}{\sqrt{3}\sigma}\right]\right\}^2}$	E_0	σ
Gamma	$\mathrm{f(E)} = \frac{\beta^\alpha \left(\frac{E-E_{min}}{E_{min}}\right)^{\alpha-1} exp\left(-\beta\frac{E-E_{min}}{E_{min}}\right)}{\Gamma(\alpha)}$	$\left(\frac{\alpha}{\beta}+1\right)E_{min}$	$\frac{\sqrt{\alpha}}{\beta}$

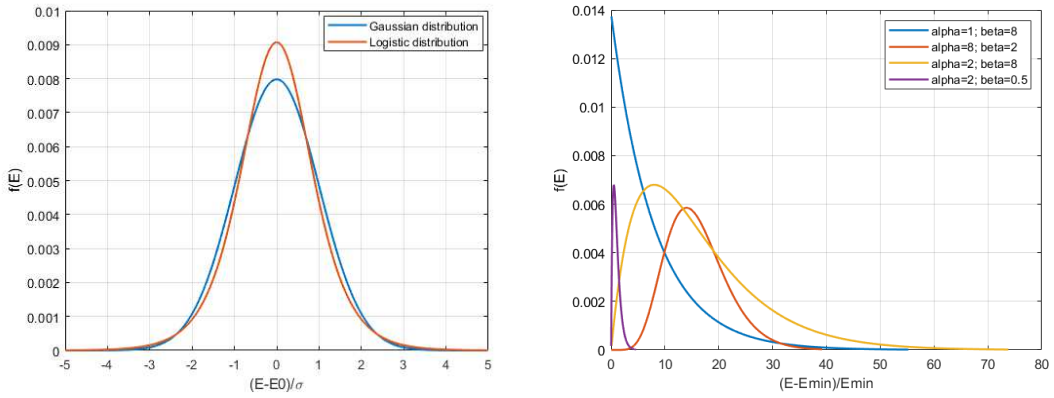


Figure 1. Examples of Gaussian, logistic distributions (left) and gamma distributions (right) with different parameters

Applying the aforementioned distributions into DAEM, the double integration (over time and energy) and the lack of an exact analytical solution required multiple precautions. Mathematical implementations were realized using the in-house MATLAB codes, including numerical discretization and integral approximation. The detailed formulation is shown hereafter.

(a) Gaussian distribution

In the discretization of Gaussian distribution, the domain $(-\infty, +\infty)$ changed to finite intervals as $[E_0 - n\sigma, E_0 + n\sigma]$, in which $n = 3$ to ensure 99.9% area coverage (figure 1). Each interval length σ was evenly divided by m to generate fine increments: $dE = \frac{\sigma}{m}$. Prior trials had determined the proper choice of increment number since it might cause solution oscillation with small values and long calculation time with large values [42]. $m = 10$ was found as good compromise for both correct representation of continuous function and concision in algorithm, the discretized activation energy for reaction i became:

$$E_i = E_0 - n\sigma + (i - 0.5)dE, \qquad i = [1, 2mn] \tag{6}$$

The discrete form of Gaussian distribution was therefore presented as:

$$f(E_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(E_i - E_0)^2}{2\sigma^2}\right) \tag{7}$$

(b) Logistic distribution

As in symmetrical distribution, finite intervals $[E_0 - 3\sigma, E_0 + 3\sigma]$ were enough for logistic distribution to ensure 99.9% area coverage. The discretization strategy of activation energy E_i was the same as equation 6, and its discretized function was expressed as

$$f(E_i) = \frac{\pi}{\sqrt{3}\sigma} \frac{exp\left[-\frac{\pi(E_i - E_0)}{\sqrt{3}\sigma}\right]}{\left\{1 + exp\left[-\frac{\pi(E_i - E_0)}{\sqrt{3}\sigma}\right]\right\}^2} \tag{8}$$

(c) Gamma distribution

Compared to common symmetric distribution, asymmetric gamma distribution had more complicated numerical implementations. The concepts of minimum (E_{min}) and maximum values (E_{max}) described the discretized activation energy. They formed a finite function

domain to avoid extremely large energies which were impossible to appear in decomposition reactions.

$$E_{max} = E_{min} \cdot \left(1 + \frac{n\alpha}{\beta}\right) \tag{9}$$

Here $n = 5$ was set to ensure a distribution function with good representation and extended in reasonable ranges. Equal partitioning of the whole interval by n_g produced fine increments dE:

$$dE = \frac{(E_{max} - E_{min})}{n_g} \tag{10}$$

in which the fine interval number n_g was set as 100 to ensure accuracy as well as short solution time. Then gamma distribution discrete normalization form was obtained:

$$f(E_i) = \frac{\beta^{\alpha} \left(\frac{E_i - E_{min}}{E_{min}}\right)^{\alpha-1} exp\left(-\beta \frac{E_i - E_{min}}{E_{min}}\right)}{\Gamma(\alpha)} \tag{11}$$

The term $\left(\frac{E_i - E_{min}}{E_{min}}\right)^{\alpha-1}$ in gamma distribution demanded preliminary determinations of discrete strategy, for instance, sudden mutations occurred when α becomes less to the unit, which produces infinite value at zero ($E_i = E_{min}$), and invalidates the numerical value at the zero point. Careful prerequisite determinations had been performed which indicated the optimal choice of the exponential distribution, a degeneracy of gamma, to facilitate model implementation. The detailed determination process could be found in supporting materials. Finally, numerical discretization in exponential distribution can be decided as:

$$E_i = E_{min} + (i - 0.5)dE, \qquad i = [1, n_g] \tag{12}$$

$$f(E_i) = \beta exp\left(-\beta \frac{E_i - E_{min}}{E_{min}}\right) \tag{13}$$

For every distribution, the characteristic time-constant of the distribution reduced with the increasing temperature level. For the smallest values of the activation energies, a simple first-order derivative might fail [43]. To avoid the difficult problem of checking the time-step for all activation energy values, the effective increment of chemical reaction dV_i was computed using the exact exponential form [44, 45]:

$$dV = \sum_{j=1}^{N_d} \sum_{i=1}^{N_p} \left[1 - exp\left(-k_i^j dt\right)\right](V_i^{j,\infty} - V_i^j(t)) \tag{14}$$

$$k_i^j = A \exp\left(-\frac{E_i^j}{RT(t)}\right); \quad V_i^{j,\infty} = V_{max}(j) \cdot m_0 \cdot f_j\left(E_i^j\right); \quad V_i^j(t=0) = 0 \tag{15}$$

in which N_d is the number of distributions, N_p is the number of increment points in each distribution. E_i^j is the ith activation energy in distribution j. $f_j\left(E_i^j\right)$ were determined only once according to the numerical implementations of every distribution at the initialization stage of the simulation. During the time-increment, the values of $V_i^j(t)$ were updated and stored in the calculation loop within the numerical domain. The updated conversion rate $X(t)$ at time t was finally obtained as follows:

$$X(t) = 1 - \frac{\sum_{j=1}^{N_d} \sum_{i=1}^{N_p} \left\{\left[1 - exp\left(-k_i^j dt\right)\right](V_i^{j,\infty} - V_i^j(t))\right\}}{m_0} \tag{16}$$

Parameter identification was based on the optimization of the objective function, *OF,* which was the residual sum of squares (RSS) between experimental and calculated conversion rates of all data points alongside the entire reaction history:

$$OF = \sum_{a=1}^{N_{exp}} \sum_{t=0}^{t_f} \left(X_{exp}(t) - X(t)\right)^2 \tag{17}$$

in which N_{exp} is the total number of experimental data and t_f is the final reaction time. Here four dynamic tests were analysed simultaneously as a learning database. This wide range of temperature-time pathways together with the large temperature range ensured quasi-complete pyrolysis was likely to provide an accurate and robust parameter determination. Regarding the different number and shape of distributions, their initial parameters were reasonably derived from the relevant literature [6, 26, 41, 46, 47] to avoid local minima in the optimization algorithm. A derivative-free method of the simplex searching algorithm was adopted, and sufficient iterations ensured successful identification by ending with no difference between the penultimate and final optimization values. The whole protocol therefore consists of the following steps: i) choosing a set of reasonable initial parameters, ii) automatic minimization algorithm, iii) perturbation of model parameters to check the robustness of the solution. If a better solution is found at stage iii), steps ii) and iii) are repeated until a stable solution is found.

To assess the effect of distribution number, a series of Gaussian-DAEMs was proposed with the distribution number ranging from one to five. Concerning the shape of distributions, three-logistic DAEM and two-Gaussian + one exponential DAEM were further proposed. Herein, only one asymmetric exponential distribution was introduced since it was enough to focus on the performances of high-temperature reactions and provided sufficient flexibility in the model [3].

To measure the discrepancy between experimental data and model simulations, two evaluation indexes were applied: root mean square error (RMSE) and the maximum deviations (D_m). They had the same meaning to DRM, aiming to provide comprehensive views on the average and maximum errors in both model identification and validation phases.

$$RMSE = \frac{\sqrt{\sum_{a=1}^{N_{exp}} \sum_{t=0}^{t_f} \left(X_{exp}(t) - X(t)\right)^2}}{N_{exp}} \quad (18)$$

326

$$D_m = Max\left|X_{exp}(t) - X(t)\right|, \ \ \forall t \in [0, t_f] \quad (19)$$

327

328 For the assessment in DTG simulations, dimensionless DTG was defined as the ratio

329 between real-time DTG signal ($\frac{dm}{dt}$) and initial anhydrous mass (m_0):

$$DTG(t) = \frac{1}{m_0}\frac{dm_t}{dt} \quad (20)$$

330 Its local residue $\left(R_L(t)\right)$ was defined by the differences between calculated ($DTG_{cal}(t)$)

331 and experimental values $DTG_{exp}(t)$:

$$R_L(t) = DTG_{cal}(t) - DTG_{exp}(t) \quad (21)$$

332 With its standard deviation of residue as:

$$\sigma_R = \sqrt{\frac{1}{t_f - 1}\sum_{t=0}^{t_f}(R_L(t) - \overline{R_L})^2} \quad (22)$$

333

334 # 4. Results and discussions

335 ## 4.1 Determination of distribution number

336 Figure 2 shows the effects of Gaussian distribution number on both identification and

337 validation stages. Initial sections of DRM curves in identifications were magnified here for

338 comparison among five models; usually, they were challenging to describe numerically

339 because of their very slow kinetics and absence of asymptotic behaviour [48]. In terms of

340 identification stages, one-Gaussian presented very poor overlaps due to its limitations in

341 describing the nature of multiple clusters of reactions [11]. Two-Gaussian demonstrated a

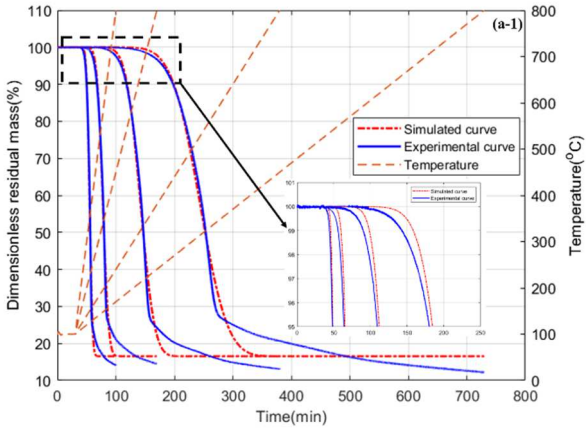
342 slight improvement, yet several noticeable errors still occurred, and the initial stages were

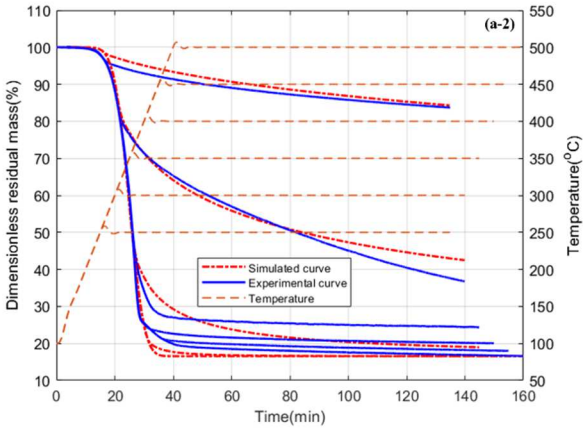
poorly produced. For the distribution number from three to five, simulations exhibited quasi-perfect agreements with all experimental data, where most overlaps indicated successful identification. The increase in distribution number represented the increase of pseudo-components, an advanced description strategy of thermal features in multi-step reactions [15]. This was further evidenced by the improved simulation accuracy in initial sections, in which error reduction could be observed by increasing distribution number, and almost complete overlap could be realized by five-Gaussian DAEM.

For the validation stages, one-Gaussian showed poor predictive ability at all temperature levels, and two-Gaussian also performed unsatisfactorily, with obvious errors from 250 to 350 °C. The three-, four- and five-Gaussian showed similarly good predictive abilities, demonstrating excellent overlap during heating periods and only slight deviations in the isothermal plateaus. Major deviations occurred on the curves of 300 °C and 350 °C, while at the other temperatures, the models showed very good predictive abilities throughout the experiments.

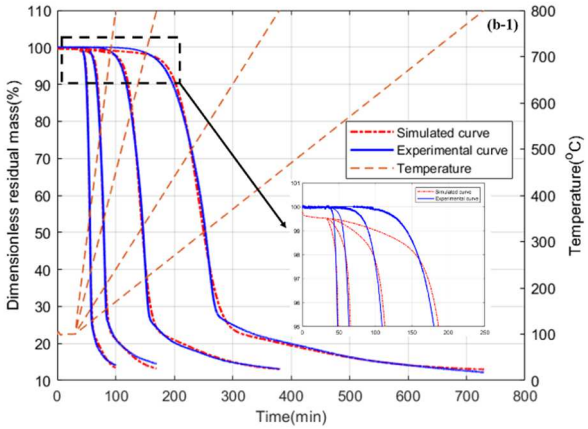
Identification **Validation**

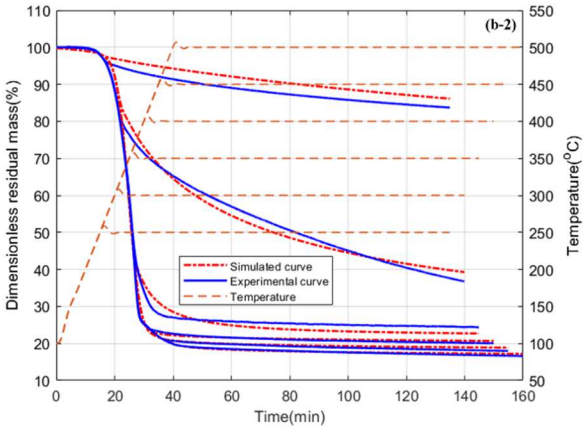
One-Gaussian DAEM



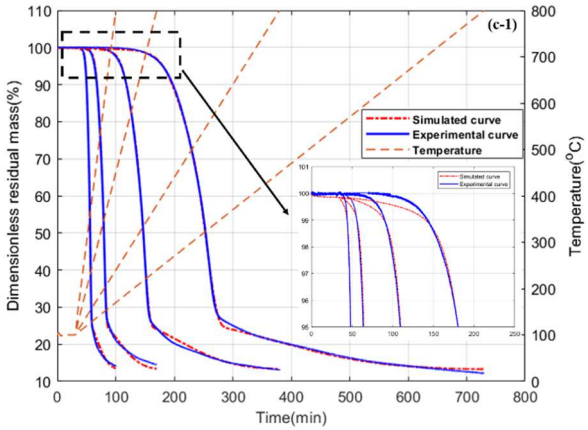


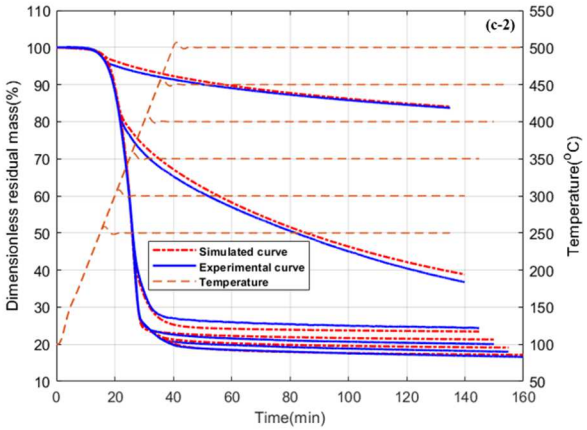
Two-Gaussian DAEM





Three-Gaussian DAEM





Four-Gaussian DAEM

Five-Gaussian DAEM

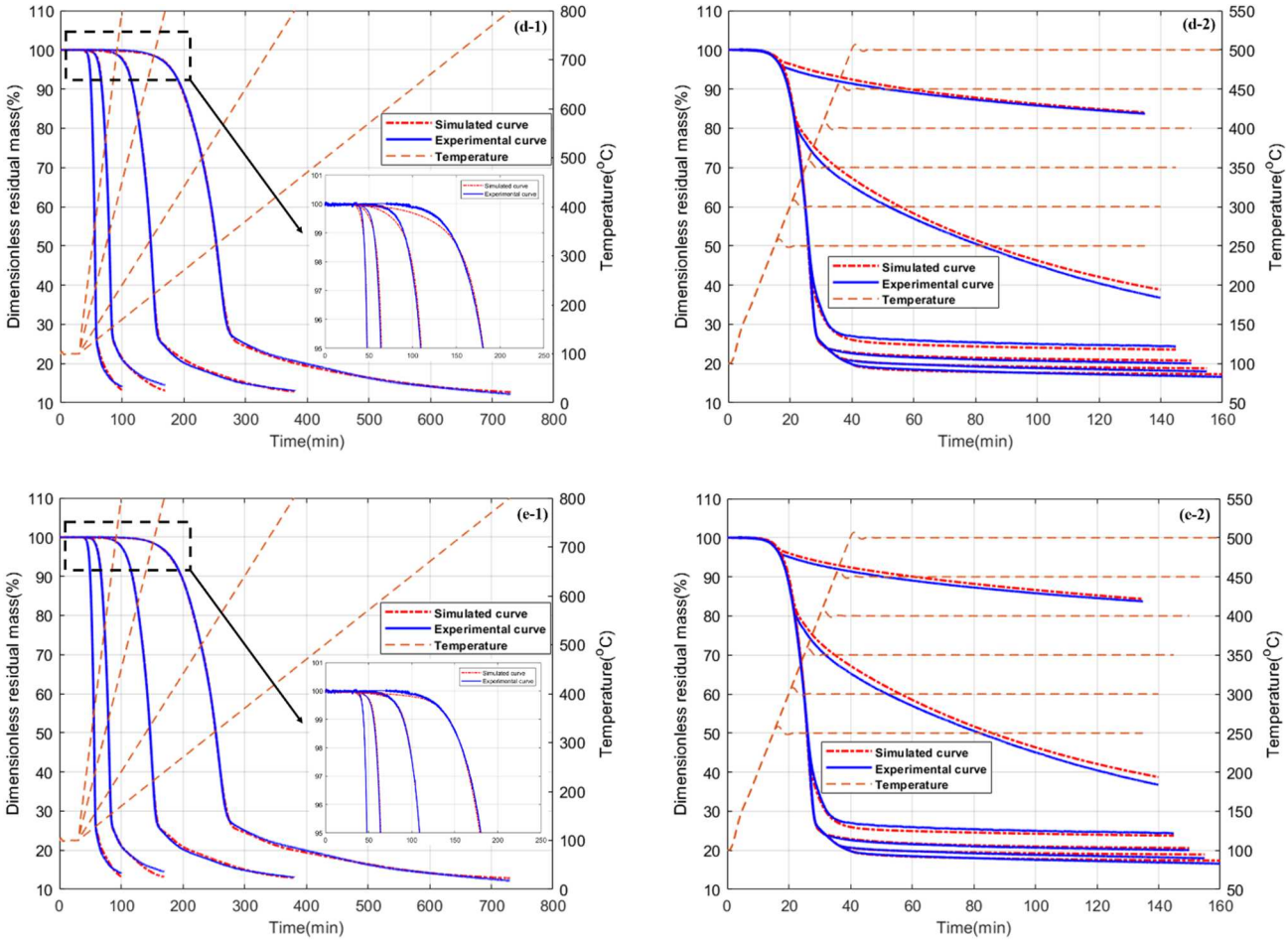


Figure 2. DRM curves of identification and validation stages and corresponding simulation curves based on identified parameters in multiple Gaussian DAEMs

(a) one-Gaussian DAEM, (b) two-Gaussian DAEM, (c) three-Gaussian DAEM, (d) four-Gaussian DAEM and (e) five-Gaussian DAEM

Identified parameters of five types of Gaussian DAEMs are listed in Table 4. The sum of weighting factors in each model ranged from 0.8345 to 0.8769, which reflected the total content of decomposable species in wood. For one-Gaussian DAEM, its coverage of activation energy was obviously incomplete due to the limitation of distribution number. The two-Gaussian DAEM introduced two distinct distributions that functioned in both low and high activation energy zones. The further increase of distribution number in three-, four- and five-Gaussian DAEM achieved more precise ranges of activation energy. Furthermore,

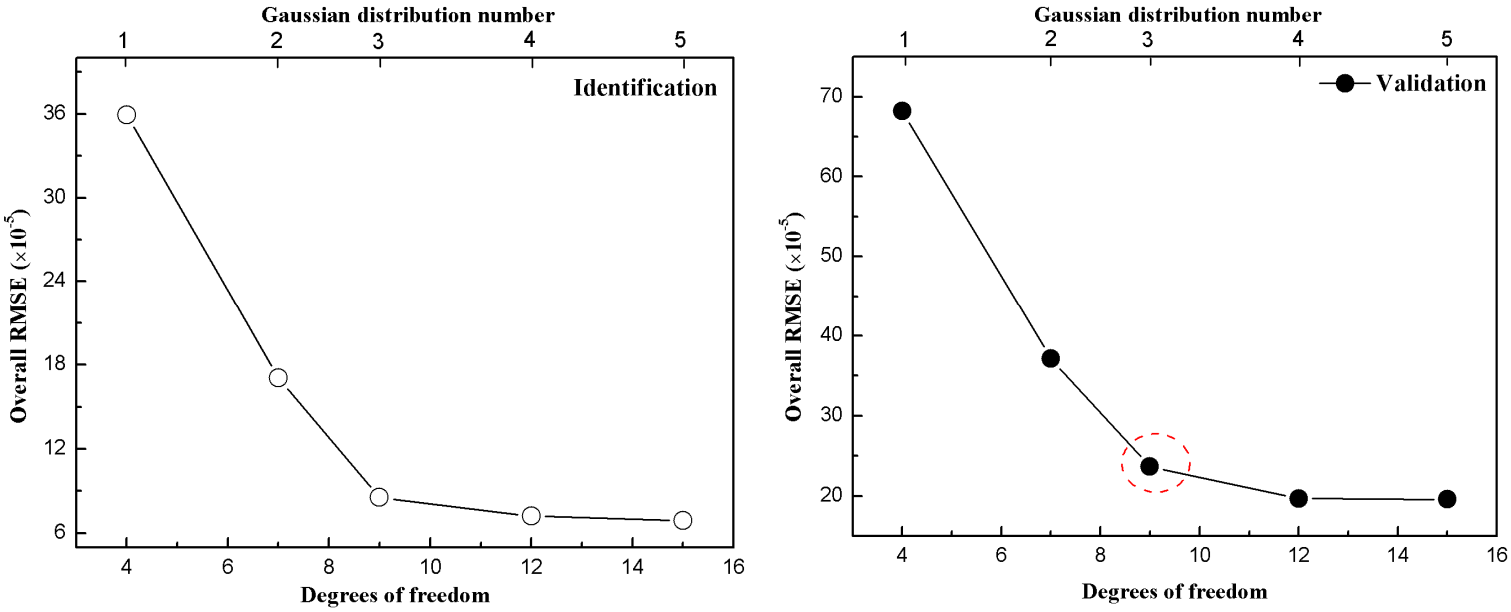
370 complicated repartitions and interactions emerged among multiple distributions. For example,
371 third and fourth Gaussian distributions were partially overlapped in four-Gaussian DAEM,
372 and the first and second distributions interacted in five-Gaussian DAEM.

373

374 Table 4. Identified model parameters of five multiple-Gaussian DAEMs

DAEM type	Distributions	V_{max}	$A(\times 10^{13} s^{-1})$	E_0 (kJ/mol)	σ (kJ/mol)
One-Gaussian	1st Gaussian	0.8345	1.56	183.78	10.59
Two-Gaussian	1st Gaussian	0.7153	0.36	175.70	6.74
	2nd Gaussian	0.1574		215.68	54.00
Three-Gaussian	1st Gaussian	0.2125	1.37	170.82	5.77
	2nd Gaussian	0.5001		185.28	1.31×10^{-9}
	3rd Gaussian	0.1553		224.67	45.80
Four-Gaussian	1st Gaussian	0.1660	1.75	170.90	4.21
	2nd Gaussian	0.4954		186.02	6.99×10^{-7}
	3rd Gaussian	0.1205		190.40	25.91
	4th Gaussian	0.0950		259.17	49.18
Five-Gaussian	1st Gaussian	0.0583	1.70	162.40	6.05
	2nd Gaussian	0.1347		172.05	5.74×10^{-3}
	3rd Gaussian	0.5284		185.87	4.01×10^{-4}
	4th Gaussian	0.0347		208.19	6.64
	5th Gaussian	0.1182		244.74	49.57

375



376 Figure 3. Relationships between the model's degrees of freedom and overall RMSE in
377 identification and validation of multiple-Gaussian-DAEM

378

379 Figure 3 shows the correlations between multiple-Gaussian DAEM's degrees of freedom
380 (D_f) and overall values of RMSE in both identification and validation stages. Followed by the
381 increase of the degrees of freedom, the overall RMSE initially decreased rapidly and then was
382 stable, indicating that prediction ability was effectively improved by increasing the
383 distribution number to three. Increasing the distribution number to four and five introduced
384 limited improvements. A high value of D_f inevitably aggravated the numerical complexity,
385 and a trade-off was, therefore, necessary with respect to the model's prediction ability and
386 complexity. Using the one-Gaussian DAEM as a reference, the decrease ratios of overall
387 residue were 45.59%, 65.40%, 71.20% and 71.36%, respectively, for two-, three-, four- and
388 five-Gaussian DAEMs during the validation stage. Using the three-Gaussian model, the
389 'inflexion point' (as highlighted with a red circle) was where a significant improvement in
390 prediction ability was gained with a relatively small increase in D_f. Even though further
391 increasing distribution number could still promote prediction potential, the disadvantage was

that one extra distribution introduced three more kinetic parameters. It was doubtless unnecessary to pursue very limited improvements at the expense of large complexity, or perhaps even worse, a decline in the model's robustness with complicated compensation effects [49]. In this context, the strategy of three Gaussian distributions was the best trade-off between model's complexity and prediction capability.

As for the identification process, proposed models showed similar matching relationships between D_f and RMSE, which also demonstrated the good trade-off of the three-Gaussian DAEM. Considering most studies have applied three-distribution DAEMs because they can correlate with main chemical constituents [5, 50], the trade-off strategy herein provides new support for using three distributions in a model.

4.2 Determination of distribution shapes

4.2.1 Three-distribution DAEMs

As a three-Gaussian model was determined to be the optimal choice among multiple-Gaussian DAEMs, its counterpart models, three-logistic and two Gaussian + one exponential DAEM, were analysed to determine the shape of the distribution (symmetry/asymmetry).

Identification and validation effects of all three-distribution DAEMs are compared in Figure 4. In the general view of identification stages for three models, as expected, simulations exhibited perfect agreements with all experimental data. Three-Gaussian and three-logistic DAEMs had similar negative deviations during the initial stages of simulations, and the two Gaussian + one exponential DAEM modified these plateaus with fewer errors. In the validation stage, the two Gaussian + one exponential DAEM showed noticeable improvements at 400, 450 and 500 °C compared to those of three-Gaussian and three-logistic DAEMs.

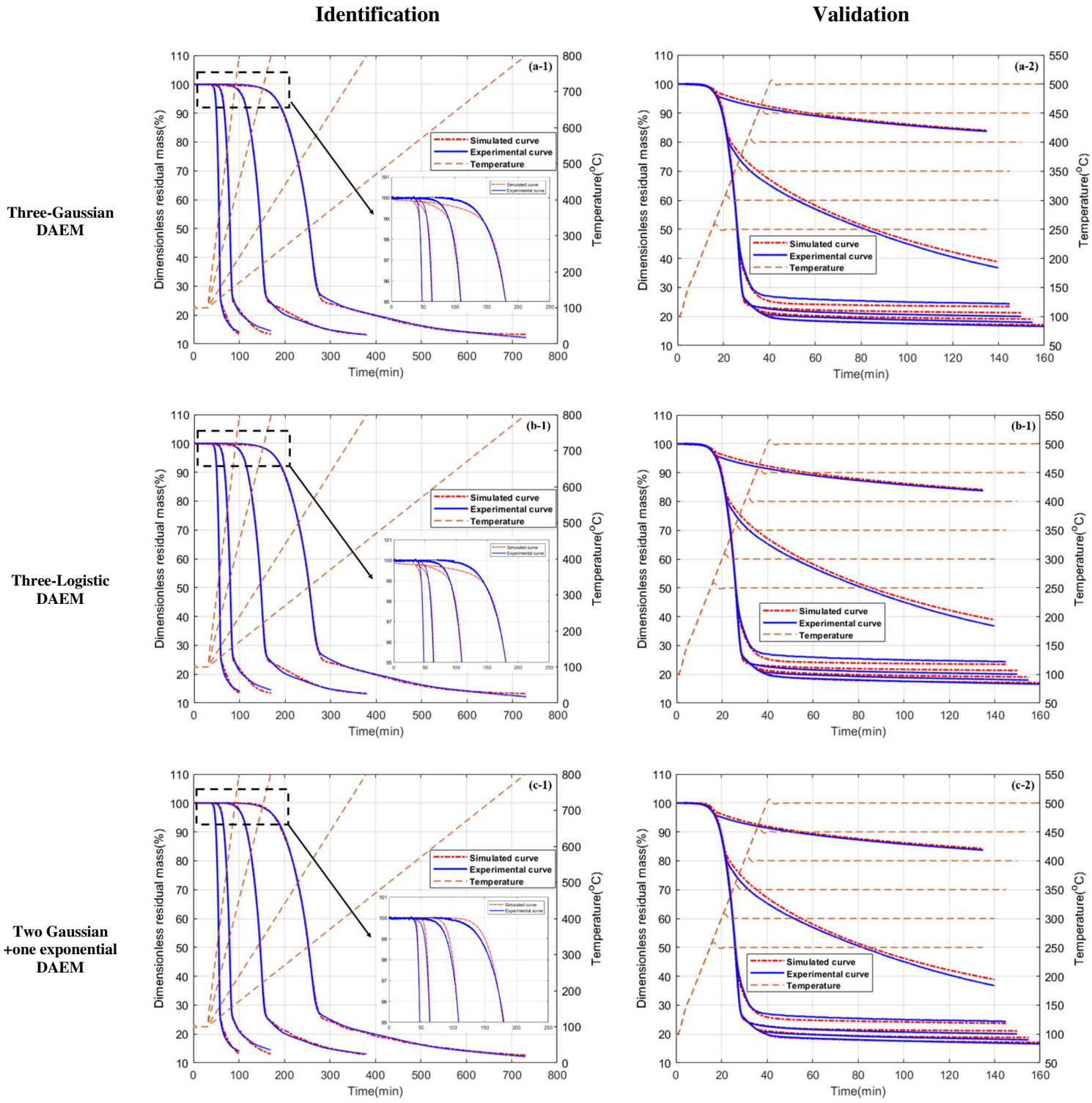


Figure 4. DRM curves of identification and validation stages and corresponding simulation curves of three types of DAEM

(a) three-Gaussian DAEM, (b) three-logistic DAEM and (c) two Gaussian+ one exponential DAEM

Identified parameters of three models are listed in Table 5. Combinations of distribution shapes made distinctive impacts on model parameters. Compared to the three-Gaussian DAEM, the three-logistic DAEM cut back its second and third weighting factors to compensate on the first. While in the two Gaussian + one exponential DAEM, the second distribution reduced its weighting factor to account for increased first and third distributions. This evident alteration was mainly caused by the third exponential distribution, which spread uniquely in the high activation energy zone due to its asymmetry, and required more weighting factors to more accurately represent chemical reactions at high temperatures. Moreover, all three models maintained their second distribution with the largest weighting factor and narrowest range; this pseudo-component could be correlated to the high content and crystalline nature of cellulose [51].

Regarding the nature of chemical collisions in reaction rate equation (eq.3), the activation energy is the barrier of relative translational motion of the reactants [52], and the pre-exponential factor represents the collision frequency that leads to successful reactions. It's therefore only meaningful to analyze the joint effect of these two parameters, which are performed in our model by fixing A as constant, and further distinguish the kinetic variation by different distribution profiles of activation energy. Herein, the excellent fitting quality in multiple experiments, which cover large temperature range and different heating rates, proves that a constant A can effectively represent pyrolysis kinetics over a large range of conditions. Thus, this average meaning of collision frequency among all reactions is meaningful and, in the meantime, keeps the model concise.

Cai [6] reviewed three-distribution DAEM in eight types of biomass, where parameter features were quite consistent with corresponding values herein. For instance, $E_{0,1}$ ranged between 169.71 kJ/mol to 186.77 kJ/mol which included our values; the narrowest ranging nature of second pseudo-component and the widest of the third were also confirmed here.

445 Várhegyi [34] tested four biomasses with the three-parallel DEAM, showing similar kinetic
446 parameters as presented in table 5, especially that $E_{0,2}$ (185 kJ/mol) almost equalled our
447 identified parameters.

448 Table 5. Identified model parameters of three-distribution DAEM

DAEM type	Distributions	V_{max}	$A(\times 10^{13} s^{-1})$	E_0 or E_{min} (kJ/mol)	σ (kJ/mol)
Three-Gaussian	1st Gaussian	0.2125	1.37	170.82	5.77
	2nd Gaussian	0.5001		185.28	1.31×10^{-9}
	3rd Gaussian	0.1553		224.67	45.80
Three-Logistic	1st Logistic	0.2233	1.43	171.25	6.63
	2nd Logistic	0.4959		185.50	2.51×10^{-6}
	3rd Logistic	0.1496		227.80	48.87
Two Gaussian + one exponential	1st Gaussian	0.2816	1.67	173.19	8.79
	2nd Gaussian	0.4199		186.15	1.66×10^{-7}
	Exponential	0.1940		175.96	α=1, β=0.43

449

450 ## 4.2.2 Overall model trade-offs

451 Table 6 and 7 list the values of RMSE and D_m among all proposed models during
452 identification and validation stages, respectively. Herein, the individual and overall values of
453 RMSE were distinguished by their different data number according to eq. 18. Since the one-
454 Gaussian and two-Gaussian DAEMs had poor prediction abilities, they will not be discussed
455 in detail hereafter. The hybrid model type (one Gaussian+ one logistic+ one exponential) was
456 previously tested, but the numerical complexity and unobvious improvements in simulation
457 quality didn't encourage us to make further investigations. But still we list its corresponding
458 data in both table 6 and 7. The main comparisons will be made among the three-distribution
459 DAEMs (three-Gaussian, three-logistic, and two-Gaussian + one exponential), and four- and
460 five-Gaussian DAEM.

461 For these models, their identification stages exhibited overall RMSE less than 9×10^{-5} and
462 the maximum deviations were reasonably small (from 1.02 to 2.74%), acting as solid proof of

successful identification. Fit qualities generally increased with a decrease in heating rates and the smallest D_m always appeared at 1 °C/min, possibly caused by the low fluctuation and stable temperature profiles at low heating rates, when thermal overshoot was less obvious [53]. The overall identification ability was in the order five-Gaussian > four-Gaussian > two Gaussian + one exponential > three-Gaussian > three-logistic. Among the three-distribution DAEMs, asymmetric exponential distribution improved identification accuracy, attributing to its capability of describing the unique behaviour at high temperatures.

Table 6: Parameter identification qualities in the dynamic tests with an average residual sum of squares and maximum deviations under each heating rate

DAEM type	RMSE ($\times 10^{-5}$)					D_m (%)			
	1°C/min	2°C/min	5°C/min	10°C/min	Overall	1°C/min	2°C/min	5°C/min	10°C/min
One-Gaussian	31.5	35.5	47.2	50.4	35.9	6.50	6.82	6.17	5.26
Two-Gaussian	12.2	17.1	27.9	33.6	17.0	4.16	4.86	4.66	4.29
Three-Gaussian	5.0	8.0	16.6	22.4	8.5	1.68	1.42	1.34	2.71
Four-Gaussian	4.0	6.6	14.4	20.7	7.2	1.21	1.56	1.44	2.59
Five-Gaussian	3.7	6.2	14.2	20.4	6.9	1.02	1.36	1.41	2.49
Three-Logistic	5.2	8.2	16.6	22.5	8.7	1.71	1.46	1.30	2.74
Gaussian + logistic + exponential	4.2	7.1	14.8	21.5	7.6	1.32	1.56	1.43	3.23
Two-Gaussian + one exponential	4.2	7.2	14.8	20.9	7.5	1.27	1.57	1.43	2.54

Table 7: Validation qualities in the static tests with an average residual sum of squares and maximum deviations under each heating rate

DAEM type	RMSE ($\times 10^{-4}$)							D_m (%)					
	250°C	300°C	350°C	400°C	450°C	500°C	Overall	250°C	300°C	350°C	400°C	450°C	500°C
One-Gaussian	3.9	6.5	9.7	9.9	6.5	4.3	6.8	2.22	5.71	5.45	5.06	5.09	5.14
Two-Gaussian	7.1	4.7	4.3	2.3	2.3	2.1	3.7	3.12	4.20	3.91	4.00	4.09	4.03
Three-Gaussian	1.8	3.9	3.4	2.3	1.9	1.0	2.4	1.29	2.20	1.99	1.46	1.54	1.51
Four-Gaussian	2.0	3.8	2.5	1.4	1.3	1.1	2.0	1.32	2.24	1.77	1.59	1.53	1.54
Five-Gaussian	2.3	3.9	2.2	1.1	1.4	1.1	2.0	1.12	2.05	1.76	1.50	1.47	1.44
Three-Logistic	1.8	3.9	3.4	2.3	1.9	1.0	2.3	1.22	2.17	2.01	1.42	1.51	1.48
Gaussian + logistic + exponential	1.6	4.1	2.4	1.7	1.3	1.1	2.0	1.22	2.34	1.84	1.73	1.68	1.69
Two-Gaussian + one exponential	1.5	3.8	2.4	1.7	1.3	1.0	1.9	1.22	2.36	1.67	1.60	1.70	1.67

In the validation of static tests, the models' true prediction potentials were therefore comparable and quantifiable at different temperature levels. The overall RMSE of five compared models ranged between 1.9×10^{-4} to 2.4×10^{-4}, which were plausibly low considering that model parameters originated from dynamic tests without any further identification. Maximum deviations were satisfactorily small in every validation stage of all models. The smallest value appeared at 250 °C (1.12% for five-Gaussian DAEM), and the largest value was at 300°C (2.36% for two Gaussian+ one exponential DAEM); such low values were considered good signs of prediction precision [12, 34]. The accuracy benefits from the comprehensive identification database and numerical training: these two indexes reflected excellent prediction potential for the five DAEM models.

Generally, the largest errors occurred at 300 °C and 350 °C among all validation phases, yet good fits were found at the other temperature levels. Because parameters were identified over the full pyrolysis condition up to 800 °C, they faced some difficulties in describing incomplete pyrolysis, even though the prediction remained rather good, and was in agreement with the literature [54] in which higher simulation errors tended to appear at low temperatures.

The mass loss under mild pyrolysis at 250 °C was slow and, in addition, quite well predicted by the DAEM model. These two reasons explain the relatively small errors found for this test.

Among these models, overall prediction ability proceeded in the order of two Gaussian + one exponential > five-Gaussian > four-Gaussian > three-Gaussian > three-logistic. Notably, the asymmetric exponential distribution exhibited excellent potential in predicting pyrolysis kinetics with plausible complexity: it provided the lowest RMSE values at 250, 300 and 500 °C, and most importantly, the lowest overall RMSE. Due to its asymmetry, exponential distribution expanded uniquely on the side of high activation energy, which only focused its accuracies in high-temperature reactions. In contrast, Gaussian and logistic distributions expanded symmetrically. When they intended to describe large ranges of high activation energies, the symmetry forced distribution to cover the same portion of low values. It conflicted with the distribution that originally represented low activation energy; thus, overall prediction potential was not effectively elevated even with additional symmetric distributions. Beyond that, the increase in distribution number ameliorated identification accuracy in the cases of four- and five-Gaussian DAEMs, yet their overall prediction potentials were not necessarily the best. It was, therefore, indispensable to conduct model identification and validation separately, to obtain correct and comprehensive assessments of the model.

The two Gaussian + one exponential DAEM demonstrated its best performances in validation stages, proving a superior strategy for the distribution's shape. This performance was further verified in Figure 5, which depicted the match relationships between qualities in both identifications and validations (overall RMSE) and model's degrees of freedom (D_f). As noticed, even though four- and five-Gaussian had lower RMSE values during model identifications, their prediction abilities were inferior to two Gaussian+ one exponential DAEM. On the other hand, three-Gaussian and three-logistic DAEM had identically small D_f as two Gaussian + one exponential DAEM, yet they both demonstrated worse prediction

potentials. Two Gaussian and one exponential successfully located both its RMSE and D_f in the low 'optimal zone' (highlighted with the red circle), which ensured accurate prediction ability while avoiding excessive numerical complexity. In this sense, two-Gaussian and one exponential provided the best trade-off between prediction ability and degrees of freedom.

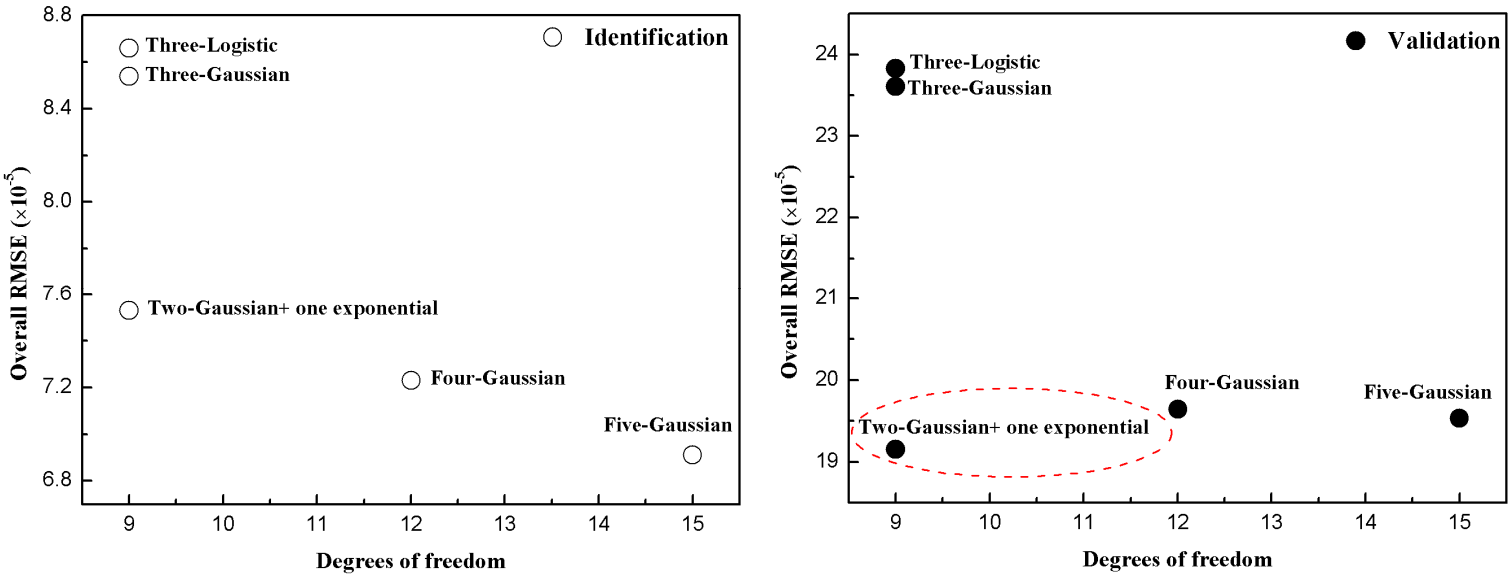


Figure 5. Degrees of freedom and overall RMSE in both identification and validation of five types of DAEM

Independent and global first-order reaction mechanism was additionally considered herein for comparison purpose. This kinetic method often worked for simple approximation that treated biomass as single or multiple pseudo-components [55, 56], in which each set of kinetic equation (Arrhenius equation) required at least two independent parameters, i.e. pre-exponential factor and global activation energy. Becidan [57] claimed 7 to 8 partial reactions were required for acceptable fit, in other words, it needed 8 to 9 independent parameters with assumption that they shared same pre-exponential factor, which was still too complicated. Its limited reaction types were unlikely appropriate descriptions of the infinite number in real pyrolysis. Worse still, the mechanism faced poor applicability in fitting multiple temperature

profiles [40, 57]. Yet in two-Gaussian + one exponential DAEM, 10 parameters were sufficient to capture essences of massive reactions, being adequately concise and meantime powerful. From this perspective, the proposed two-Gaussian + one exponential again presented as excellent model choice for kinetic determinations.

To further confirm the superiority of two-Gaussian + one exponential DAEM, three-n*th*-order mechanism was also applied for comparison [58]. Herein, pre-exponential factor (A) was assumed identical in three reactions, and each scheme had its specific activation energy (E) and reaction order (n). The superposition of three n*th*-order reaction schemes was subsequently fitted in both dynamic and static tests for parameter identification and verification (table 8). Regarding the overall RMSE in identification, three-n*th*-order model had much larger deviation than in two-Gaussian + one exponential DAEM (RMSE=7.6× 10^{-5}), revealing worse fitting quality. More importantly, this model had worse prediction capacity compared to two-Gaussian + one exponential DAEM (RMSE=2.0 × 10^{-4} in validation), yet its numerical complexity was hardly simplified ($D_f = 10$), which again supported the better performances of proposed DAEM.

Table 8. Identified parameters of three-n*th*-order model and overall RMSE in identification and validation stages

Pseudo-component	V_{max}	A (s^{-1})	E (kJ/mol)	n	Overall RMSE in identification	Overall RMSE in validation
First	0.3234		152.88	0.83		
Second	0.4728	2.98× 10^{11}	165.53	1.10	1.9× 10^{-4}	4.5 × 10^{-4}
Third	0.2931		227.95	0.69		

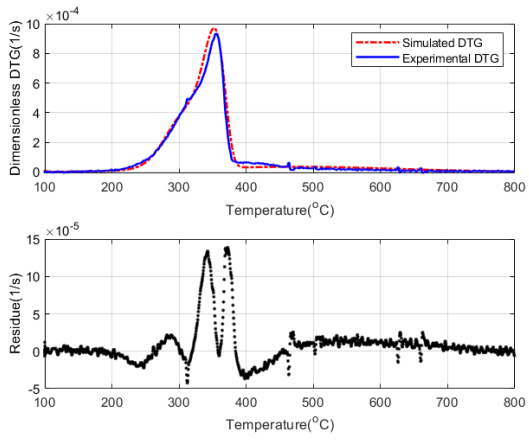
4.3 DTG simulation

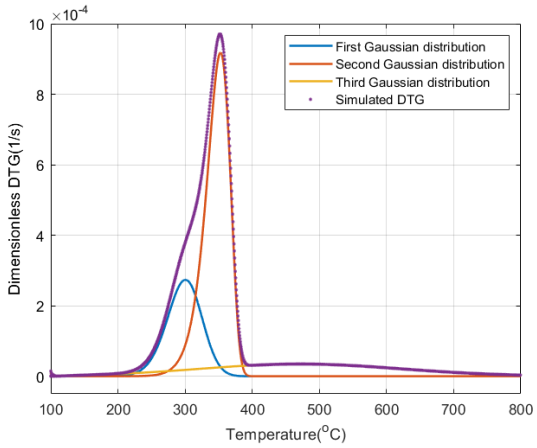
DTG simulations were conducted to assess the performance of different distributions in the five comparable DAEMs. Figure 6 illustrated DTG simulations by the overall model and individual distributions at 5 °C/min. The heating rate was chosen due to its moderate baseline fluctuations; DTG simulations at other heating rates are within the supplementary material. It

was obvious that the overall DTG simulations of five models overlapped substantially with experimental data, and the only visible deviations occurred near the DTG peak. For analysis, the main variations of local residue $\left(R_L(t)\right)$ were manually divided into three noticeable sections here, ranging between 100–300, 300–400 and 400–800 °C, respectively. The first and third sections moderately fluctuated, and hemicellulose decomposition was mainly attributed to the first. Meanwhile, lignin decomposition and secondary reactions were responsible for the third section [59]. The second section showed the most severe fluctuations with two obvious peaks; they were believed to originate from the mixed pyrolysis of hemicellulose and cellulose [60]. Cellulose decomposed rapidly within a very narrow temperature range after its crystallites melted [61] and it partially merged with hemicellulose. Inevitably more detectable errors emerged in this section during simulations.

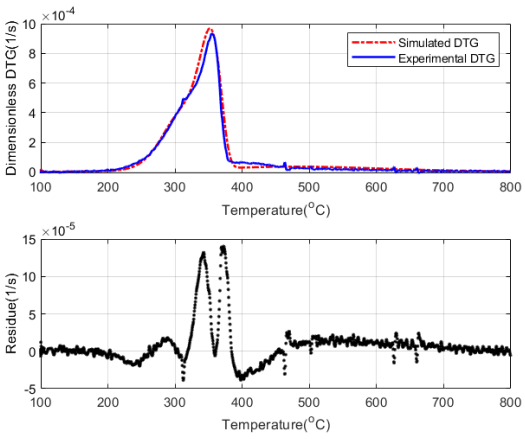
Overall DTG with local residues | **Separate and overall DTG**

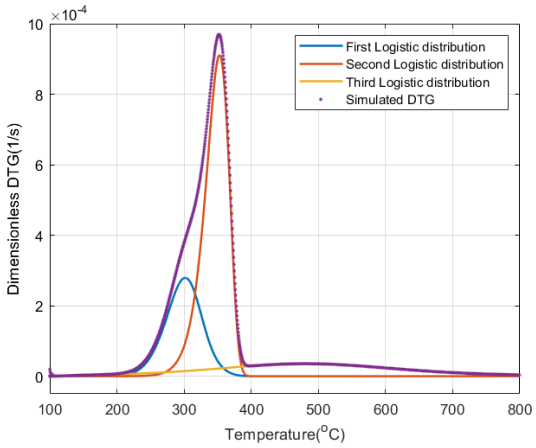
Three-Gaussian DAEM



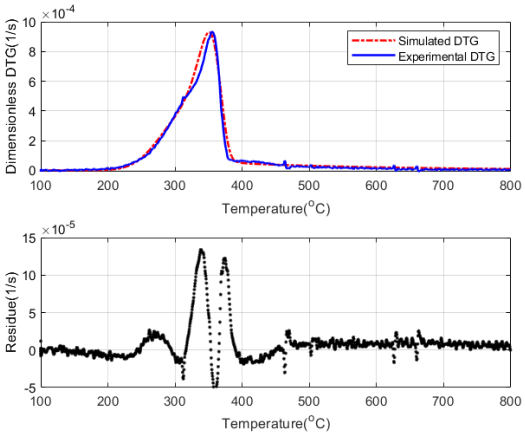


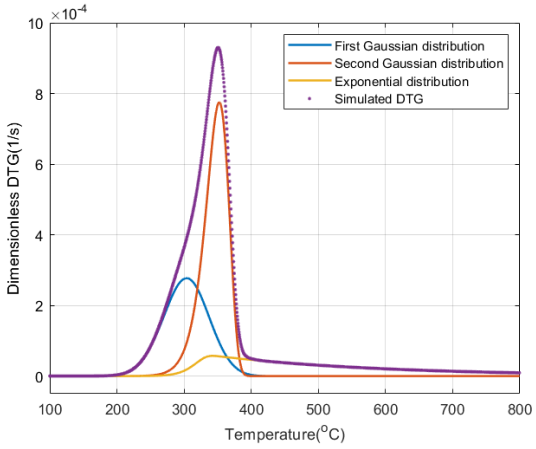
Three-Logistic DAEM



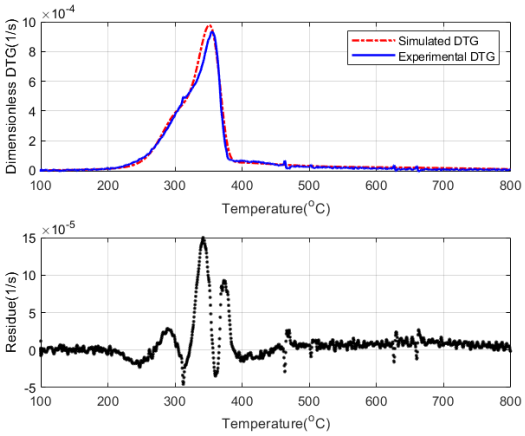


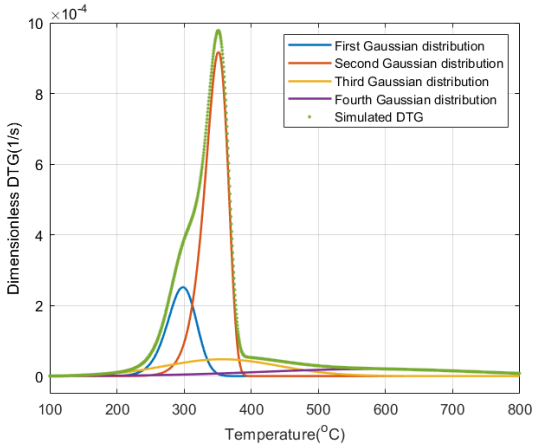
Two Gaussian +one exponential DAEM





Four-Gaussian DAEM





Five-Gaussian DAEM

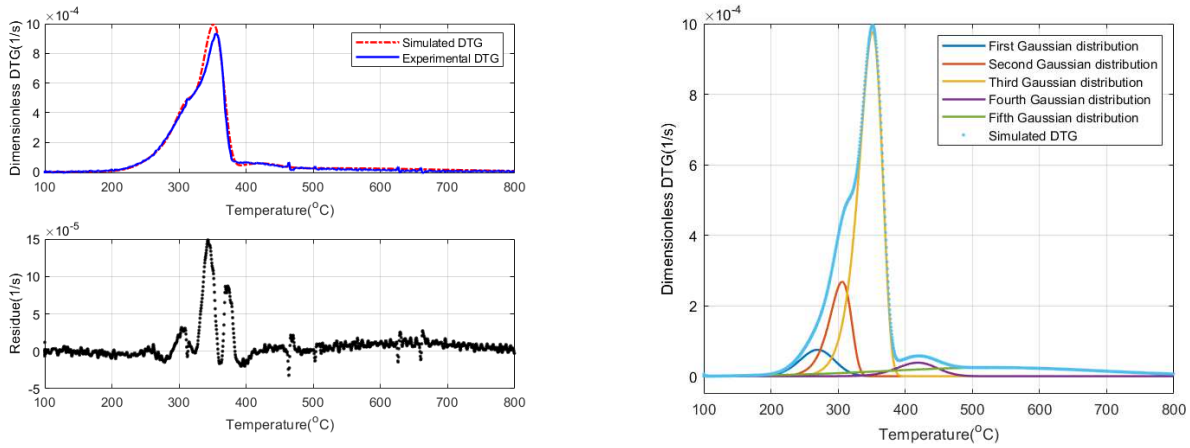


Figure 6. Experimental, simulated DTG curves and corresponding local residues (left), separate and overall DTG simulations (right) of five DAEM at the heating rate of 5 °C/min

Standard deviations (σ_R) of sectional simulations were compared in Table 9. Among the DAEMs with three distributions, two Gaussian + one exponential DAEM exhibited the best simulation improvements, especially in the third section wherein the asymmetric exponential uniquely functioned. Further compared to two Gaussian + one exponential DAEM, four-Gaussian DAEM showed only slight improvements in first and second sections, and five-Gaussian DAEM improved in all three sections. However, as previously emphasized, it was not practical nor feasible to increase numerical complexity for limited error improvement. Using the two Gaussian + one exponential DAEM as a reference, the error reduction ratios in three sections realized by four-Gaussian DAEM were 4.34 %, 1.60 %, 0 %, respectively, and 42.03 %, 10.49 %, 8.33 %, respectively, by five-Gaussian DAEM. The five-Gaussian DAEM provided the only notable improvement in the first section, which represented a very small part of pyrolysis kinetics. So, here the asymmetric exponential distribution again presented a good trade-off strategy to improve accuracy and maintain numerical concision.

Table 9. The sectional standard deviation of DTG residue at the heating rate of 5°C/min of five models

Temperature section (°C)	σ_R ($\times 10^{-6}$)				
	Three-Gaussian	Three-Logistic	Two Gaussian+one exponential	Four-Gaussian	Five-Gaussian
100–300	7.2	7.4	6.9	6.3	4.0
300–400	57.3	57.3	56.2	55.3	49.5
400–800	11.9	12.2	7.2	7.2	6.6

Separate DTG simulation further depicted each pseudo component's contribution to the overall decomposition rate (right row of Figure 6). The three-Gaussian and three-logistic DAEMs had same decomposition ranges of three pseudo-components: 200–375 °C (first pseudo-component), 250–400 °C (second pseudo-component) and 100–800 °C (third pseudo-component). Compared to these two models, two Gaussian + one exponential DAEM altered the decomposition range of first pseudo-component (200–400 °C), induced no change for second pseudo-component (250–400 °C), and increased initial decomposition temperature of the third pseudo-component (270–800 °C). For the four-Gaussian DAEM, four pseudo-components decomposed in the temperature ranges as 220–360, 250–400, 160–550 and 100–800 °C respectively. Finally, the decomposition temperature ranges of five-Gaussian DAEM's pseudo-components were: 165–375, 200–350, 240–400, 280–520 and 100–800 °C, respectively.

It was always worthwhile to identify pseudo-components as the signatures of major chemical components. The totality of specific species could be described by one unique distribution, and further assist in understanding the multi-component mechanism of biomass pyrolysis. In this vein, Table 10 shows relevant thermogravimetric decomposition temperatures of the three main constituents in biomass. Generally, hemicellulose consists of polysaccharides that are heterogeneously branched and are non-covalently bonded to the

surface of cellulose microfibril [62]. The less stable chemical structures make hemicellulose more reactive with lower temperature ranges of decomposition. Cellulose is a saturated linear polysaccharide with high polymerization and degree of crystallinity and decomposes in very narrow temperature ranges [63]. Lignin is a set of irregular phenolic polymers consisting of more than four substituted phenyl propane, which endow lignin with large decomposition ranges during the entire pyrolysis process [64]. Many published values provided reliable references for featuring pseudo-components. By comparing chemical constituents' decomposition temperatures with model calculations, it was, therefore, feasible to correspond the three pseudo-components to hemicellulose, cellulose and lignin respectively in the three-Gaussian, three-logistic and two Gaussian + one exponential DAEM, which were well-accepted strategies [6, 34, 65]. For the four-Gaussian DAEM, first and second pseudo-components were linked with hemicellulose and cellulose, while the third and fourth pseudo-components possibly presented the two-stage decomposition scheme of lignin [66]. In the five -Gaussian DAEM, the first and second pseudo-components together represent hemicellulose, then the third pseudo-component was assigned to cellulose. The fourth and fifth components were attributed to two overlapped partial reactions, the scission of oxygen functional groups and rearrangement of the carbon skeleton were believed responsible for these two distributions, respectively [48].

639 Table 10. Summary of individual biomass component's decomposition temperature range

Heating rate (°C/min)	Decomposition temperature range (°C)			Reference
	Hemicellulose	Cellulose	lignin	
10	220–315	315–400	>400	[67]
10	200–320	280–360	140–600	[68]
20	200–350	260–430	200–500	[69]
20	253–308	319–368	259–482	[70]
20	244–324	294–371	197–653	[71]
50	250–350	350–500	>500	[72]
60	200–327	327–450	200–550	[2]
<5	225–325	324–375	250–500	[73]

640

641 Peak temperatures observed in both experimental and simulated DTG at different heating
642 rates are listed in Table 11. The peak temperatures corresponded to the extreme reaction stage
643 of cellulose decomposition, and they followed a decreasing trend with an increase of heating
644 rate, which was caused by the altered pyrolysis kinetics relating to inter-particle heat transfer
645 [74]. In general, five models reproduced the peak temperatures accordingly at four heating
646 rate conditions, with considerably small errors less than 5 °C. It was clear that multiple-
647 distribution DAEMs were capable of distinguishing the effects of heating rates, which has
648 rarely been observed in relevant researches. Peak values were usually applied in Kissinger
649 procedure for determining activation energy [75], as a direct and easy method. However, its
650 limitations arose in merely staying at observation stage based on existing experimental data.
651 By comparison, multiple-distribution DAEMs herein presented a precise description of peak
652 temperatures and confirmed the prediction abilities already proved on mass loss. Therefore, it
653 would be safe to conclude that the proposed models could advance one step further as
654 prediction tools in isoconversional kinetics. DAEM may be applied for the accurate
655 description of global pyrolysis kinetics under different reaction conditions.

656

657 Table 11. Peak decomposition temperature from experiments and model simulations

Heating rate (°C/min)	Experimental	Three-Gaussian	Three-Logistic	Two Gaussian+ one exponential	Four-Gaussian	Five-Gaussian
1	366.58	363.58	363.58	362.59	362.59	362.59
2	354.73	351.73	352.23	350.73	350.73	351.23
5	339.06	337.08	337.08	335.88	336.08	336.48
10	329.44	326.64	326.64	325.13	325.43	325.83

658

659 ## Conclusion

660 A rigorous strategy of both identification and validation was applied in the DAEM for
661 analysing the effects of distribution number and shape. First, in a series of Gaussian-DAEMs
662 with distribution number ranging from one to five, three-distribution was determined as the
663 best trade-off between prediction ability and degrees of freedom. Logistic and exponential
664 distributions were proposed to account for distribution shapes. Exponential distribution
665 allowed good simulations for high-temperature reactions, and together with two Gaussian
666 distribution, they exhibited as the best strategy in terms of both prediction capacity and
667 numerical concision. The overall DTG simulation at 5 °C/min was analysed within three
668 major variation sections, where two-Gaussian + one exponential DAEM achieved obvious
669 error reduction with plausible numerical concision. Degradation temperature ranges by
670 separate DTG simulations evidenced the correspondences between pseudo-components and
671 chemical components, and DAEM's was able to distinguish the effect of heating rate on the
672 peak decomposition temperature.

673 ## Acknowledgements

674 The authors are grateful for the financial support of the Conseil Général de la Marne, Grand
675 Reims and the Région Grand Est, France. The support from the China Scholarship Council
676 (CSC) is also gratefully acknowledged.

677

Reference

[1] M. Balat, M. Balat, E. Kırtay, H. Balat, Main routes for the thermo-conversion of biomass into fuels and chemicals. Part 1: Pyrolysis systems, Energy conversion and Management 50(12) (2009) 3147-3157.

[2] Q. Liu, Z. Zhong, S. Wang, Z. Luo, Interactions of biomass components during pyrolysis: A TG-FTIR study, Journal of Analytical and Applied Pyrolysis 90(2) (2011) 213-218.

[3] A.K. Burnham, R.L. Braun, Global kinetic analysis of complex materials, Energy & Fuels 13(1) (1999) 1-22.

[4] A. Anca-Couce, Reaction mechanisms and multi-scale modelling of lignocellulosic biomass pyrolysis, Progress in Energy and Combustion Science 53 (2016) 41-79.

[5] C. Di Blasi, Modeling chemical and physical processes of wood and biomass pyrolysis, Progress in energy and combustion science 34(1) (2008) 47-90.

[6] J. Cai, W. Wu, R. Liu, An overview of distributed activation energy model and its application in the pyrolysis of lignocellulosic biomass, Renewable and Sustainable Energy Reviews 36 (2014) 236-246.

[7] G. Pitt, The kinetic of the evolution of volatile products from coal, Fuel 41 (1962) 267-274.

[8] J.L. Hillier, T.H. Fletcher, Pyrolysis kinetics of a Green River oil shale using a pressurized TGA, Energy & fuels 25(1) (2011) 232-239.

[9] A. Soria-Verdugo, N. Garcia-Hernando, L. Garcia-Gutierrez, U. Ruiz-Rivas, Analysis of biomass and sewage sludge devolatilization using the distributed activation energy model, Energy Conversion and Management 65 (2013) 239-244.

[10] Q. Xiong, J. Zhang, F. Xu, G. Wiggins, C.S. Daw, Coupling DAEM and CFD for simulating biomass fast pyrolysis in fluidized beds, Journal of Analytical and Applied Pyrolysis 117 (2016) 176-181.

[11] H. Liu, M.S. Ahmad, H. Alhumade, A. Elkamel, R.J. Cattolica, Three pseudo-components kinetic modeling and nonlinear dynamic optimization of Rhus Typhina pyrolysis with the distributed activation energy model, Applied Thermal Engineering 157 (2019) 113633.

[12] J. Cai, W. Wu, R. Liu, Sensitivity analysis of three-parallel-DAEM-reaction model for describing rice straw pyrolysis, Bioresource technology 132 (2013) 423-426.

[13] G. Várhegyi, H. Chen, S. Godoy, Thermal decomposition of wheat, oat, barley, and Brassica carinata straws. A kinetic study, Energy & fuels 23(2) (2009) 646-652.

[14] J. Wu, Y. Liao, Y. Lin, Y. Tian, X. Ma, Study on thermal decomposition kinetics model of sewage sludge and wheat based on multi distributed activation energy, Energy 185 (2019) 795-803.

[15] J. Zhang, T. Chen, J. Wu, J. Wu, Multi-Gaussian-DAEM-reaction model for thermal decompositions of cellulose, hemicellulose and lignin: Comparison of N2 and CO2 atmosphere, Bioresource Technology 166 (2014) 87-95.

[16] A.K. Burnham, R.L. Braun, H.R. Gregg, A.M. Samoun, Comparison of methods for measuring kerogen pyrolysis rates and fitting kinetic parameters, Energy & Fuels 1(6) (1987) 452-458.

[17] K. Miura, A new and simple method to estimate f (E) and k0 (E) in the distributed activation energy model from three sets of experimental data, Energy & Fuels 9(2) (1995) 302-307.

[18] D. Xu, M. Chai, Z. Dong, M.M. Rahman, X. Yu, J. Cai, Kinetic compensation effect in logistic distributed activation energy model for lignocellulosic biomass pyrolysis, Bioresource technology 265 (2018) 139-145.

[19] C.C. Lakshmanan, N. White, A new distributed activation energy model using Weibull distribution for the representation of complex kinetics, Energy & fuels 8(6) (1994) 1158-1167.

[20] M. Li, L. Liu, L. Jiang, F.-H. Gou, J.-H. Sun, Application of distributed activation energy models to polymer pyrolysis: Effects of distributed model selection, characteristics, validation, and sensitivity analysis, Fuel 254 (2019) 115594.

[21] T. Xu, F. Xu, Z. Hu, Z. Chen, B. Xiao, Non-isothermal kinetics of biomass-pyrolysis-derived-tar (BPDT) thermal decomposition via thermogravimetric analysis, Energy Conversion and Management 138 (2017) 452-460.

[22] A. Dhaundiyal, S.B. Singh, J.C. Salcedo-Reyes, Asymptotic approximations to the isothermal pyrolysis of deodara leaves using gamma distribution, Universitas Scientiarum 22(3) (2017) 263-284.

[23] B.P. Boudreau, B.R. Ruddick, On a reactive continuum representation of organic matter diagenesis, American Journal of Science 291(5) (1991) 507-538.
[24] K. Balakrishnan, Exponential distribution: theory, methods and applications, Routledge2018.
[25] S. Nadarajah, S. Kotz, The beta exponential distribution, Reliability engineering & system safety 91(6) (2006) 689-697.
[26] B. de Caprariis, M.L. Santarelli, M. Scarsella, C. Herce, N. Verdone, P. De Filippis, Kinetic analysis of biomass pyrolysis using a double distributed activation energy model, Journal of Thermal Analysis and Calorimetry 121(3) (2015) 1403-1410.
[27] Y. Lin, Z. Chen, M. Dai, S. Fang, Y. Liao, Z. Yu, X. Ma, Co-pyrolysis kinetics of sewage sludge and bagasse using multiple normal distributed activation energy model (M-DAEM), Bioresource technology 259 (2018) 173-180.
[28] C. Chen, W. Miao, C. Zhou, H. Wu, Thermogravimetric pyrolysis kinetics of bamboo waste via Asymmetric Double Sigmoidal (Asym2sig) function deconvolution, Bioresource technology 225 (2017) 48-57.
[29] Z. Dong, Y. Yang, W. Cai, Y. He, M. Chai, B. Liu, X. Yu, S.W. Banks, X. Zhang, A.V. Bridgwater, Theoretical analysis of double Logistic distributed activation energy model for thermal decomposition kinetics of solid fuels, Industrial & Engineering Chemistry Research 57(23) (2018) 7817-7825.
[30] A. Soria-Verdugo, E. Goos, N. García-Hernando, Effect of the number of TGA curves employed on the biomass pyrolysis kinetics results obtained using the Distributed Activation Energy Model, Fuel Processing Technology 134 (2015) 360-371.
[31] J. Cai, L. Ji, Pattern search method for determination of DAEM kinetic parameters from nonisothermal TGA data of biomass, Journal of Mathematical Chemistry 42(3) (2007) 547-553.
[32] K. Santos, F. Lobato, T. Lira, V. Murata, M.A. Barrozo, Sensitivity analysis applied to independent parallel reaction model for pyrolysis of bagasse, Chemical Engineering Research and Design 90(11) (2012) 1989-1996.
[33] A.I. Ferreiro, M. Rabaçal, M. Costa, A combined genetic algorithm and least squares fitting procedure for the estimation of the kinetic parameters of the pyrolysis of agricultural residues, Energy Conversion and Management 125 (2016) 290-300.
[34] G. Várhegyi, B. Bobály, E. Jakab, H. Chen, Thermogravimetric study of biomass pyrolysis kinetics. A distributed activation energy model with prediction tests, Energy & Fuels 25(1) (2011) 24-32.
[35] S. Scott, J. Dennis, J. Davidson, A. Hayhurst, An algorithm for determining the kinetics of devolatilisation of complex solid fuels from thermogravimetric experiments, Chemical engineering science 61(8) (2006) 2339-2348.
[36] Y. Lin, Y. Tian, Y. Xia, S. Fang, Y. Liao, Z. Yu, X. Ma, General distributed activation energy model (G-DAEM) on co-pyrolysis kinetics of bagasse and sewage sludge, Bioresource technology 273 (2019) 545-555.
[37] M.S. Ahmad, H. Liu, H. Alhumade, M.H. Tahir, G. Çakman, A. Yıldız, S. Ceylan, A. Elkamel, B. Shen, A modified DAEM: To study the bioenergy potential of invasive Staghorn Sumac through pyrolysis, ANN, TGA, kinetic modeling, FTIR and GC–MS analysis, Energy Conversion and Management 221 (2020) 113173.
[38] C.N. Arenas, M.V. Navarro, J.D. Martínez, Pyrolysis kinetics of biomass wastes using isoconversional methods and the distributed activation energy model, Bioresource technology 288 (2019) 121485.
[39] A. Soria-Verdugo, M. Rubio-Rubio, E. Goos, U. Riedel, Combining the lumped capacitance method and the simplified distributed activation energy model to describe the pyrolysis of thermally small biomass particles, Energy Conversion and Management 175 (2018) 164-172.
[40] S. Vyazovkin, A.K. Burnham, J.M. Criado, L.A. Pérez-Maqueda, C. Popescu, N. Sbirrazzuoli, ICTAC Kinetics Committee recommendations for performing kinetic computations on thermal analysis data, Thermochimica acta 520(1-2) (2011) 1-19.
[41] S. Cavagnol, J.F. Roesler, E. Sanz, W. Nastoll, P. Lu, P. Perré, Exothermicity in wood torrefaction and its impact on product mass yields: From micro to pilot scale, The Canadian Journal of Chemical Engineering 93(2) (2015) 331-339.

[42] M. Güneş, S. Güneş, The influences of various parameters on the numerical solution of nonisothermal DAEM equation, Thermochimica acta 336(1-2) (1999) 93-96.
[43] I. Turner, P. Rousset, R. Rémond, P. Perré, An experimental and theoretical investigation of the thermal treatment of wood (Fagus sylvatica L.) in the range 200–260 C, International Journal of Heat and Mass Transfer 53(4) (2010) 715-725.
[44] P. Perré, I.W. Turner, A 3-D version of TransPore: a comprehensive heat and mass transfer computational model for simulating the drying of porous media, International Journal of heat and mass transfer 42(24) (1999) 4501-4521.
[45] R. Remond, I. Turner, P. Perre, Modeling the drying and heat treatment of lignocellulosic biomass: 2D effects due to the product anisotropy, Drying Technology 28(8) (2010) 1013-1022.
[46] R.K. Mishra, K. Mohanty, Pyrolysis kinetics and thermal behavior of waste sawdust biomass using thermogravimetric analysis, Bioresource technology 251 (2018) 63-74.
[47] J. Zhang, T. Chen, J. Wu, J. Wu, A novel Gaussian-DAEM-reaction model for the pyrolysis of cellulose, hemicellulose and lignin, Rsc Advances 4(34) (2014) 17513-17520.
[48] G. Varhegyi, M.J. Antal Jr, E. Jakab, P. Szabó, Kinetic modeling of biomass pyrolysis, Journal of analytical and Applied Pyrolysis 42(1) (1997) 73-87.
[49] P.J. Barrie, The mathematical origins of the kinetic compensation effect: 1. The effect of random experimental errors, Physical Chemistry Chemical Physics 14(1) (2012) 318-326.
[50] Z. Chen, M. Hu, X. Zhu, D. Guo, S. Liu, Z. Hu, B. Xiao, J. Wang, M. Laghari, Characteristics and kinetic study on pyrolysis of five lignocellulosic biomass via thermogravimetric analysis, Bioresource Technology 192 (2015) 441-450.
[51] L.-q. Jiang, A.-q. Zheng, J.-g. Meng, X.-b. Wang, Z.-l. Zhao, H.-b. Li, A comparative investigation of fast pyrolysis with enzymatic hydrolysis for fermentable sugars production from cellulose, Bioresource technology 274 (2019) 281-286.
[52] M. Menzinger, R. Wolfgang, The meaning and use of the Arrhenius activation energy, Angewandte Chemie International Edition in English 8(6) (1969) 438-444.
[53] Z. Ma, J. Xie, N. Gao, C. Quan, Pyrolysis behaviors of oilfield sludge based on Py-GC/MS and DAEM kinetics analysis, Journal of the Energy Institute 92(4) (2019) 1053-1063.
[54] A. Soria-Verdugo, L. Garcia-Gutierrez, L. Blanco-Cano, N. Garcia-Hernando, U. Ruiz-Rivas, Evaluating the accuracy of the distributed activation energy model for biomass devolatilization curves obtained at high heating rates, Energy conversion and management 86 (2014) 1045-1049.
[55] G. Varhegyi, M.J. Antal Jr, T. Szekely, P. Szabo, Kinetics of the thermal decomposition of cellulose, hemicellulose, and sugarcane bagasse, Energy & fuels 3(3) (1989) 329-335.
[56] J.J. Manya, E. Velo, L. Puigjaner, Kinetics of biomass pyrolysis: a reformulated three-parallel-reactions model, Industrial & engineering chemistry research 42(3) (2003) 434-441.
[57] M. Becidan, G. Várhegyi, J.E. Hustad, Ø. Skreiberg, Thermal decomposition of biomass wastes. A kinetic study, Industrial & engineering chemistry research 46(8) (2007) 2428-2437.
[58] T.R. Rao, A. Sharma, Pyrolysis rates of biomass materials, Energy 23(11) (1998) 973-978.
[59] H. Yang, R. Yan, H. Chen, D.H. Lee, C. Zheng, Characteristics of hemicellulose, cellulose and lignin pyrolysis, Fuel 86(12-13) (2007) 1781-1788.
[60] T. Kan, V. Strezov, T.J. Evans, Lignocellulosic biomass pyrolysis: A review of product properties and effects of pyrolysis parameters, Renewable and Sustainable Energy Reviews 57 (2016) 1126-1140.
[61] M. Weinstetn, A. Broido, Pyrolysis-crystallinity relationships in cellulose, Combustion Science and Technology 1(4) (1970) 287-292.
[62] P. McKendry, Energy production from biomass (part 1): overview of biomass, Bioresource technology 83(1) (2002) 37-46.
[63] S. Prasad, A. Singh, H. Joshi, Ethanol as an alternative fuel from agricultural, industrial and urban residues, Resources, Conservation and Recycling 50(1) (2007) 1-39.
[64] D. Mohan, C.U. Pittman Jr, P.H. Steele, Pyrolysis of wood/biomass for bio-oil: a critical review, Energy & fuels 20(3) (2006) 848-889.
[65] T. Chen, J. Zhang, J. Wu, Kinetic and energy production analysis of pyrolysis of lignocellulosic biomass using a three-parallel Gaussian reaction model, Bioresource Technology 211 (2016) 502-508.
[66] H. Chen, N. Liu, W. Fan, Two-step consecutive reaction model and kinetic parameters relevant to the decomposition of Chinese forest fuels, Journal of Applied Polymer Science 102(1) (2006) 571-576.

[67] H. Yang, R. Yan, H. Chen, C. Zheng, D.H. Lee, D.T. Liang, In-depth investigation of biomass pyrolysis based on three major components: hemicellulose, cellulose and lignin, Energy & Fuels 20(1) (2006) 388-393.
[68] S.D. Stefanidis, K.G. Kalogiannis, E.F. Iliopoulou, C.M. Michailof, P.A. Pilavachi, A.A. Lappas, A study of lignocellulosic biomass pyrolysis via the pyrolysis of cellulose, hemicellulose and lignin, Journal of analytical and applied pyrolysis 105 (2014) 143-150.
[69] S. Wang, X. Guo, K. Wang, Z. Luo, Influence of the interaction of components on the pyrolysis behavior of biomass, Journal of Analytical and Applied Pyrolysis 91(1) (2011) 183-189.
[70] E. Biagini, F. Barontini, L. Tognotti, Devolatilization of biomass fuels and biomass components studied by TG/FTIR technique, Industrial & Engineering Chemistry Research 45(13) (2006) 4486-4493.
[71] E. Biagini, L. Tognotti, A generalized procedure for the devolatilization of biomass fuels based on the chemical components, Energy & fuels 28(1) (2014) 614-623.
[72] K. Raveendran, A. Ganesh, K.C. Khilar, Pyrolysis characteristics of biomass and biomass components, Fuel 75(8) (1996) 987-998.
[73] F. Shafizadeh, Pyrolytic reactions and products of biomass, Fundamentals of thermochemical biomass conversion, Springer1985, pp. 183-217.
[74] M. Stenseng, A. Jensen, K. Dam-Johansen, Investigation of biomass pyrolysis by thermogravimetric analysis and differential scanning calorimetry, Journal of analytical and applied pyrolysis 58 (2001) 765-780.
[75] H.E. Kissinger, Reaction kinetics in differential thermal analysis, Analytical chemistry 29(11) (1957) 1702-1706.