Prediction of instantaneous yield of bio-oil in fluidized biomass pyrolysis using long short-term memory network based on computational fluid dynamics data

Hanbin Zhong [a,b], Zhenyu Wei [a], Yi Man [c], Shaowei Pan[d], Juntao Zhang [a], Ben Niu [a,b], Xi Yu[e], Yi Ouyang[f], Qingang Xiong [c, *]

[a] *Xi'an Key Laboratory of Low-carbon Utilization for High-carbon Resources, Xi'an Shiyou University, Xi'an, Shaanxi, 710065, China*

[b] *Shaanxi Engineering Research Center of Green Low-carbon Energy Materials and Processes, Xi'an Shiyou University, Xi'an, Shaanxi, 710065, China*

[c] *State Key Laboratory of Pulp and Paper Engineering, South China University of Technology, Guangzhou, 510641, China*

[d] *School of Computer Science, Xi'an Shiyou University, Xi'an, Shaanxi, 710065, China*

[e] *College of Engineering and Physical Sciences, Aston University, Birmingham, B4 7ET, United Kingdom*

[f] *Laboratory for Chemical Technology, Ghent University, Ghent, 9052, Belgium*

Abstract Computational fluid dynamics (CFD) is an effective tool to investigate biomass fast pyrolysis in fluidized bed reactor for bio-oil production, while it requires huge computational time when optimizing operating conditions or simulating large/industrial units. Machine learning (ML) is a promising approach to achieving both accuracy and efficiency. In this work, a reduced-order model including long short-term memory (LSTM) layer, pooling layer, and fully connected layer

* Corresponding author. Email: qingangxiong@scut.edu.cn (Q. Xiong)

was established to predict future mass flow rates by training the historical CFD data. Unsteady mass flow rates, which are normally used to determine product yields, were treated as data of time series in ML. Influencing factors, such as sequence length, number of neurons, learning rate, subsequences order (shuffle or not), number of LSTM layers, and ratio of testing set, were evaluated to obtain their optimal values. The developed LSTM model framework and training process showed good applicability for the dataset of different species and temperatures. Product yields predicted by the derived LSTM were in good agreement with those obtained by CFD, while nearly 30% computational effort was saved. Thus, it is clearly seen that the well-predicted fluctuating characteristics and final product yields are helpful to improve accuracy of process simulation for digitalizing key reactors and building smart factories.

Keywords*:* Bio-oil; Biomass fast pyrolysis; Fluidized bed; CFD; Machine learning; LSTM

Word Count: 6508

Nomenclature

Abbreviations

Adam	Adaptive moment estimation
ANN	Artificial neural network
BP	Back propagation
CFD	Computational fluid dynamics
CNN	Convolutional neural network
CPFD	Computational particle fluid dynamics
DEM	Discrete element method

DL	Deep learning
FC	Fully connected
LSTM	Long short-term memory
MAPE	Mean absolute percentage error
MFM	Multi-fluid model
ML	Machine learning
MSE	Mean square error
RMSE	Root mean square error
RNN	Recurrent neural network

Symbols

c_t	Value of cell state at time *t*
$\tilde{c}_t$	Value of candidate memory unit at time *t*
b	Bias vector
f_t	Value of forget gate at time *t*
h_t	Value of hidden state at time *t*
i_t	Value of input gate at time *t*
o_t	Value of output gate at time *t*
s	Value of real flowrates, kg/s
tanh	Hyperbolic tangent function
U, W	Weight matrices

X	Sequence data
x	Training data
$\hat{y}$	Predicted value
y	Actual value

Greek symbols

σ	Sigmoid function

Subscripts

c	Cell state
f	Forget gate
i	Input gate
max	Maximum
min	Minimum
nor	Normalization
o	Output gate
t	Time, s

1. Introduction

With the rapid depletion of conventional fossil fuels and the urgent challenge of global warming, renewable biomass resource has become one of the most important ways to realize sustainability (Wu et al., 2022). Among the available methods for biomass utilization, fast pyrolysis is a promising technology due to its ability to convert abundant raw biomass to transportable liquid bio-oil for down-stream large-scale processing, which is especially useful for regions with highly distributed biomass resources (Fakayode et al., 2023). Fluidized bed is an efficient type of reactor for biomass fast pyrolysis because of its high mass and heat transfer rates, and has been developed by many researchers in the past few decades. However, the experimental study of fluidized bed for biomass fast pyrolysis normally costs lots of time and money due to the complex gas-solid hydrodynamics and pyrolysis reaction mechanism.

On the other hand, computational fluid dynamics (CFD) has increasingly become an indispensable tool to facilitate the advancement of the fluidized bed reactor for biomass fast pyrolysis (Xiong et al., 2017). For example, Houston et al. (2022) used CFD to track individual particles in a bubbling fluidized bed reactor for biomass fast pyrolysis, and the influence of operating conditions on product yields was studied. Lu et al. (2022) summarized the advanced sub-models coupled with MFiX multi-scale CFD solver to model biomass fast pyrolysis. These sub-models were fully validated against experiments and can be applied to different feedstocks. Lao et al. (2022) evaluated four different models for describing the effects of intra-particle transport phenomena and secondary tar cracking. It was found that the non-isothermal model with tar cracking was the best, which was applied to investigate the effects of particle size, inflow rate of

external heat, and operation temperature. What's more important, CFD simulation is also considered as one of the most important tools to digitalize key units to build smart factory for Industry 4.0 (Aversano et al., 2021; Silvestri, 2021). In addition, as pointed out by (Vikram et al., 2021), results of process simulation using ideal or simplified hydrodynamics commonly do not coordinate well with actual operating conditions. While CFD predictions are more accurate than those of process simulation, CFD has been proved to be very helpful for improving accuracy of process simulation (Arora et al., 2017; Porrazzo et al., 2016), which is useful for many applications except building smart factories, such as techno-economic and environmental assessments (Shoaib Ahmed Khan et al., 2022).

Nevertheless, it is very time-consuming to optimize operating parameters in a wide range or simulate large/industrial fluidized bed reactor using CFD due to its huge computational load. In recent years, with the fast development of computational theory and ability, it's worth noting that machine learning (ML) or deep learning (DL) has become a promising method to accelerate CFD simulation in the field of chemical engineering (Zhu et al., 2022). For instance, Zhu et al. (2020) developed ML models based on highly-resolved CFD simulations to account for sub-gird corrections in coarse-grid CFD simulations, and the prediction is in good accordance with experimental results. Ouyang et al. (2022a) designed a novel hybrid method to switch CFD and artificial neural network (ANN) model automatically according to deviation rate, saving up to 40% computation time. Lu et al. (2021) utilized a convolutional neural network (CNN) to accelerate the computational speed of discrete element method (DEM) using a novel multiscale loss function, which significantly reduced model fluctuations caused by training steps. Ouyang et al. (2022b)

conducted a data-driven ANN model and invoked its parameters in ANSYS Fluent for furtherly compiling. The implementation results demonstrated that it could speed up turbulent reactive flow simulation by an order of magnitude with respect to using a Lagrangian probability density function (PDF) approach. Ladický et al. (2015) developed a ML-based regression forest method for estimating discrete particle movement in smooth particle hydrodynamic simulations with large time step, which could quickly approximate next-frame position and velocity of the current particle. Kochkov et al. (2021) proposed an end-to-end DL method to improve approximation for modeling two-dimensional turbulent flows, which kept the same finer resolution as baseline solvers but resulted in 40- to 80-fold computational speedups. Specially for biomass fast pyrolysis in fluidized-bed reactors, Zhong et al. (2020) developed back-propagation (BP) ANNs based on CFD data from multi-fluid model (MFM) simulation. Time-averaged species distributions in the reactor at different temperatures were predicted with much less computational time but good accuracy. Kim et al. (2022) trained eight ML models based on computational particle fluid dynamics (CPFD) simulation data. Reaction temperature and gas residence time were used as inputs, and the predicted product yields were in good agreement with CPFD results. Thus, in summary, if based on CFD data, ML is a promising approach to improve process simulation within feasible amounts of time.

One of the most interested indicators of reactor performance for biomass fast pyrolysis is the product yield of bio-oil. In CFD simulations, product yield of bio-oil is normally calculated based on flow rates of species at reactor outlet, which is temporally unsteady due to the intense gas-solid hydrodynamics in fluidized beds. From the point view of ML, these unsteady flow rates can be treated as data of time series. For data of time series, ML methods such as long short-term memory

(LSTM) network and CNN can be used to predict future values after training the historical data. Forecasting models based on the data of time series have been widely applied in other areas to address prediction problems related to time series. For example, Sun and Huang (2020) proposed a novel hybrid carbon-price prediction method combining secondary decomposition algorithm with BP neural network. In the case analysis of Hubei, Beijing, and Shanghai market, this designed model could predict carbon prices accurately. Zhao et al. (2019) constructed a LSTM-FC neural network for PM2.5 concentration prediction, which revealed a better predictive performance compared to ANN and LSTM models on the same dataset. Wang et al. (2020b) established an earthquake prediction model with two-dimension input based on LSTM network, adding the decomposition method at the same time. Numerous testing and contrast experiments revealed that this model could well capture spatiotemporal correlations and be capable of warning in the area without sensors and monitors.

However, to the best of our knowledge, there is almost no report on ML of data of time series from CFD of fluidized-bed biomass fast pyrolysis. For example, Xie et al. (2022) reported that LSTM and CNN-LSTM models can reduce 90~95% computational time compared with CFD-DEM simulations while keeping good accuracy for predicting maximum height and mixing index in a bi-disperse solid-liquid fluidized bed. Bazai et al. (2021) also showed that an encoder-decoder CNN based on CFD data of time series can quickly predict instantaneous contours of particle volume fraction at next time steps in a pseudeo-2d fluidized bed.

Therefore, in order to reduce computational requirements of CFD of biomass fast pyrolysis in fluidized bed reactors, the LSTM method was used to predict future bio-oil flow rates by training

the historical CFD data. Six main influencing factors, including sequence length, number of neurons, learning rate, subsequences order (shuffle or not), number of LSTM layers, and ratio of testing set, were investigated to choose optimal parameters. Then, the optimized LSTM model was applied to predict future flow rates of other species, i.e., gas, biomass, and char, to verify its applicability. In addition, CFD data from different pyrolysis temperatures was also tested. Product yields obtained by LSTM were in good agreement with CFD predictions, while much less computational effort was needed, showing its great potential to facilitate process simulation for digitalizing key reactors and building smart factories.

2. Dataset from CFD simulation

In CFD simulation of fluidized bed biomass pyrolysis, the instantaneous mass flow rates at outlet can be obtained as time series dataset for ML. Many methods such as MFM and CFD-DEM have been successfully applied to reveal the flow and reaction characteristics of biomass fast pyrolysis in fluidized bed reactors (Xiong et al., 2017). This work used MFM due to its relatively low requirements of computational effort. The model details, simulation conditions, and main results are given as follows.

2.1. CFD model

The MFM based on Eulerian-Eulerian method used in this study has been reported in our previous studies (Zhong et al., 2020; Zhong et al., 2019), as shown in Table S1. There are three phases, including one gas phase and two solid phases for biomass and sand particles, respectively. The kinetic granular theory, which assumes that the random motion of particles is analogous to the motion of molecules in a gas, is used to close the governing equations for each solid phase (Ding

and Gidaspow, 1990). All the phases are formulated through a set of conservation equations for mass, momentum, energy, species, and granular temperature. The gas-solid and solid-solid drag coefficient was accounted by the Gidaspow and Syamlal drag model, respectively (Huilin and Gidaspow, 2003; Syamlal, 1987). The improved Shafizadeh-Chin scheme considering secondary bio-oil cracking was used to describe the reaction mechanism of biomass fast pyrolysis. The particle shrinkage model developed based on the mass conservation at the particle scale was applied to determine the real-time particle size during simulation, which provides the opportunity to successfully capture the entrainment behavior of the reacted biomass particles (Zhong et al., 2016a; Zhong et al., 2020). However, the intra-particle heat conduction effect was neglected in this work due to its weak influence on the product yields of relatively small biomass particles (Zhong et al., 2019).

2.2. Simulation conditions

The fast pyrolysis of red oak in a bubbling fluidized bed (Xue et al., 2012) was simulated in two dimensions (2D) as shown in Fig. S1. The physical properties and basic simulation conditions are the same as our previous work, as shown in Table S2 (Zhong et al., 2020; Zhong et al., 2019). The initial packed sand particles were 55 mm high with a solid volume fraction of 0.59. Biomass particles entered into the reactor with a diameter of 325 μm at 300 K, and the bed temperature was maintained at 773 K. The structured computational grids with cells of 3.81 mm (width) × 3.65 mm (height) were implemented, and a time step of 0.001 s was used in simulation according to our previous investigations on the effects of mesh size and time step (Zhong et al., 2016a; Zhong et al., 2016b). In order to avoid the numerical errors at start-up, biomass inlet was switched on after 10 s

simulation with only sand particles, and total simulation time was 100 s.

2.3. CFD results

After CFD simulation, datasets of instantaneous mass flow rates for different species at reactor outlet are obtained as shown in Fig. 1, which will be used to train LSTM models. Because these data are recorded for every 0.01 s, the data number is 9000 for each species. In order to validate the CFD approach, product yields of CFD simulation are obtained by integrating the mass flow rates at outlet over the last 30 s. The product yield of char is determined by combing biomass and char species at outlet to be consistent with the experiments which weigh particles collected by cyclones. The predicted product yields are compared with experimental results as shown in Table 1. Clearly, CFD simulation results are in good agreement with experimental data, which indicates that it is reasonable to use these CFD datasets for ML in the following sections. It should be mentioned that the poor agreement of gas yield may be due to the mass imbalance in the experiments. Because the mass of bio-oil and char can be determined more accurately, it is reasonable to deduce that the mass imbalance was mainly caused by the weighing error of gas product. Therefore, the experimental gas yield may be 15.3 wt% after minus a mass imbalance of 5.2 wt%, which is also close to the CFD predicted result (13.73 wt%). In order to provide different CFD datasets, a CFD case at 723 K was also performed as shown in Fig. S2.

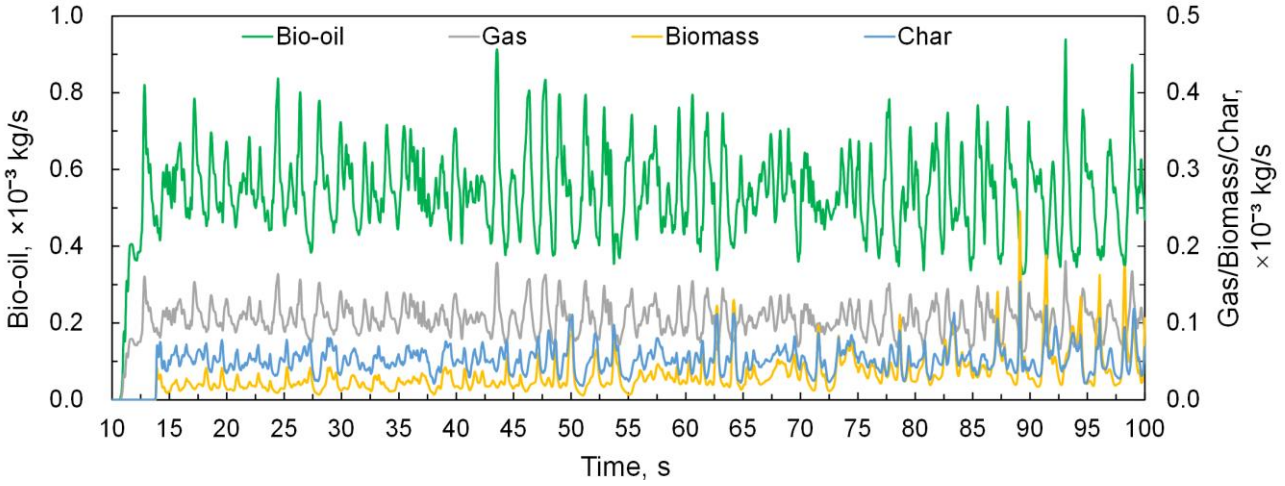

Fig. 1. CFD predicted mass flow rates at 773 K.

Table 1 Experimental and CFD predicted product yields at 773 K.

Product yield, wt%	Bio-oil	Gas	Char	Residual
Experiment (Xue et al., 2012)	71.7±1.4	20.5±1.3	13.0±1.5	-
CFD	71.39	13.73	13.74	1.14

3. LSTM model development

A brief introduction of LSTM is given in this section. The methods for data processing before ML are described. The model framework mainly contains LSTM layer, pooling layer, and fully connected layer. After model training, the mass flow rates in testing time range can be predicted by a rolling prediction method, and the model's performance is evaluated by the standards based on errors.

3.1. LSTM network

Recurrent neural network (RNN) is a suitable network for solving the problems involving time series analysis, since it is able to store the information of historical data and use them in the prediction of future value. However, traditional RNN usually performs poorly due to the frequently

occurred problems of gradient explosion and disappearance in capturing long-term correlations. While as an advanced form of RNN, LSTM network can effectively avoid these problems by introducing input gate (i_t), forget gate (f_t) and output gate (o_t) as shown in Fig. 2. The detailed calculation process of LSTM network (Sagheer and Kotb, 2019; Wang. et al., 2020a) can be divided into three parts:

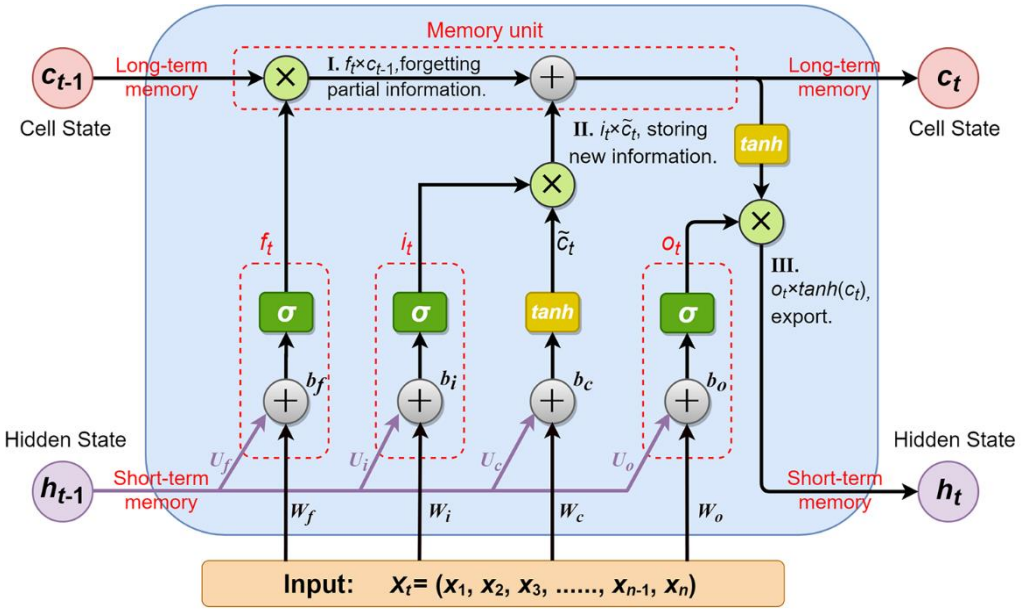

Fig. 2. LSTM network structure.

Firstly, forget gate (f_t) is calculated according to input data (X_t) at time t and previous hidden state (h_{t-1}), shown in Eq. (1). Then it is controlled between 0 and 1 by sigmoid function (i.e., σ in Fig. 2), which determines what old information should be discarded in previous cell state (c_{t-1}). A value closing to 0 means lots of information from c_{t-1} are removed, while all of them are remained if f_t equals 1.

$$f_t = \sigma\left(W_f X_t + U_f h_{t-1} + b_f\right) \tag{1}$$

Where, W_f and U_f are the weight matrices related to forget gate, b_f is the corresponding bias vector of it.

Secondly, it needs to be decided which information from X_t is going to be stored in cell state. Input gate (i_t) is designed as Eq. (2), and its similar processing with forget gate (f_t) keeps its range within (0, 1) for filtering. The new candidate unit ($\tilde{c}_t$) is created as Eq. (3), which is an activation of input data at present time and hidden state (h_{t-1}). Then, new cell state (c_t) is updated by combining partial storing information from X_t and old remaining information in forgetting processing, shown in Eq. (4).

$$i_t = \sigma\left(W_i X_t + U_i h_{t-1} + b_i\right) \tag{2}$$

$$\tilde{c}_t = \tanh\left(W_c X_t + U_c h_{t-1} + b_c\right) \tag{3}$$

$$c_t = f_t \times c_{t-1} + i_t \times \tilde{c}_t \tag{4}$$

Where, tanh represents hyperbolic tangent function in the range of (-1, 1). W_i and U_i are the weight matrices of input gate (i_t), while W_c and U_c are related to candidate unit state ($\tilde{c}_t$). b_i and b_c are the bias vectors of input gate and candidate unit state, respectively.

Finally, new hidden state (h_t) can be obtained by output gate (o_t) and cell state (c_t). The calculation formulas are as follows:

$$o_t = \sigma\left(W_o X_t + U_o h_{t-1} + b_o\right) \tag{5}$$

$$h_t = o_t \times \tanh\left(c_t\right) \tag{6}$$

Where, W_o and U_o are the weight matrices of output gate (o_t), and b_o is the bias vector of output gate.

3.2. Model training

The flowchart of model training process is shown in Fig. 3, mainly including following parts:

(1) Data pre-processing

Although mass flow rates data from 10s to 100s are generated by CFD in section 2, not all the data are used for model training and testing. The sharp increasing mass flow rates at initial stage (10 s~15 s) as shown in Fig. 1 are discarded due to their negative impact on model training. After importing data, these CFD data is divided into two groups, i.e., training data and testing data, according to the ratio of testing set as shown in Table 3. Then, numerous subsequences with a certain length are extracted from training data according to the sequence length and number of predicting steps. For instance, if sequence length is 10 and a subsequence starts from 20.00 s aiming to predict the mass flow rate of the future 0.01 s (one step), the data of this subsequence is between 20.00 and 20.10 s with a subsequence length of 11. In addition, before splitting into training set and validation set, the order of subsequences is kept the same as when extracting or randomly shuffled. Lastly, the training set and validation set are normalized to achieve better accuracy, which is described by Eq (7). And the predicted instantaneous mass flow rates are calculated according to the anti-normalization function as shown in Eq (8).

$$s_{\mathrm{nor}} = \frac{s_t - x_{\min}}{x_{\max} - x_{\min}} \tag{7}$$

$$\hat{y}_t = \hat{y}_{\mathrm{nor}} \times (x_{\max} - x_{\min}) + x_{\min} \tag{8}$$

Where, s_t is the real mass flowrate at time t. $x_{\min}$ and $x_{\max}$ is the minimum and maximum value of training data, respectively. s_{nor} and $\hat{y}_{\mathrm{nor}}$ are normalized real and predicted data between (0,1), respectively. $\hat{y}_t$ is the predicted instantaneous mass flow rate.

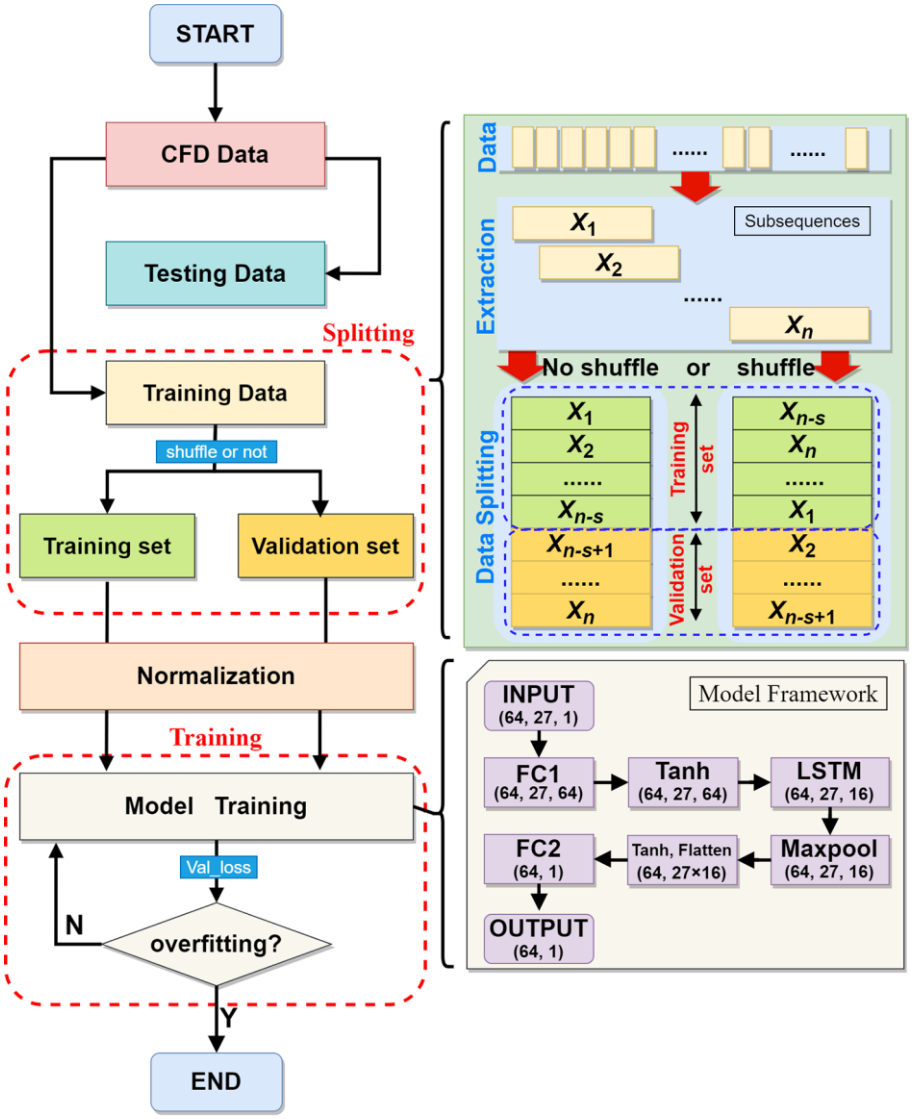

Fig. 3. Flowchart of model training process.

(2) Model Framework

The model framework mainly includes LSTM layer, pooling layer, fully connected layer, and activation function. The input/output dimensionalities are marked for each layer in the model framework part of Fig. 3. One-dimension sequence data with a length of 27 are imported for training, and the batch size is set to 64. The first fully connected layer (FC1) with 64 neurons extends dimensions to capture the relationships between input data, while the second fully connected layer (FC2) with one neuron combines characteristic information to receive ultimate

predictions of batch sequence data. The pooling layer not only integrates local feature information, but also prevents overfitting to a certain extent. The LSTM with 16 neurons, as an essential role in the prediction of time series data, realizes forecasting ability by capturing long-term correlation of sequences. The activation function is used to strengthen model's ability of nonlinear expression, while the flatten function reduces feature dimensionality to 2D level.

(3) Model Configuration

Table 2 summarized the fixed model parameters used in the training process. The maximum of training epochs is set to 500. In order to reach a faster speed and better reproducibility, batch size, the subset size of training sample (e.g., 10 out of 100), is assigned to 64, and the global network random seed of this network is set to 115, based on the results of a preliminary study. Besides, mean square error (MSE) is used as the loss function in the model training process. Adaptive moment estimation (Adam) is employed to optimize model since it is computationally efficient and well-suited for problems that are large in terms of data and/or parameters. In addition, the code for early termination is added during model training in order to prevent overfitting, which is realized by stopping the training process when validation loss reaches nearly constant. The model training process was carried out on PaddlePaddle platform of Baidu with Intel Core i9-12900K.

Table 2 Basic model configuration.

Parameter	Epoch	Batch size	SEED	Optimizer	Loss function	Early Stopping
Configuration	500	64	115	Adam	MSE	True

(4) Influencing factors

Six influencing factors, including sequence length, number of neurons, learning rate,

subsequences order (shuffle or not), number of LSTM layers, and ratio of testing set as shown in Table 3 were investigated using the dataset of bio-oil to choose the optimal parameter group. Thereinto, suitable sequence length of subsequences ensures that the proposed model has a preliminary roll-prediction ability. Magnitude of learning rate determines the size of weight-iteration step during model training, which affects the speed of model convergence. Training and validation subsequences were shuffled before splitting if the shuffle option was set. The trained models with and without shuffle were compared to draw a conclusion which one is qualified for the forecasting job. Normally, it is necessary to adjust the number of neurons and layers of LSTM to improve forecasting capacity. Ratio of testing set represents the proportion of testing set in the entire dataset. For instance, the ratio of 15/85 means the data for testing is the last 15 s in the whole dataset (85 s). Because a larger ratio of testing set means less CFD computation time, case 6 was thereby done to minimize CFD simulation time without losing the prediction accuracy of LSTM.

Table 3 Detailed model and training parameters of cases 1-6.

Parameter	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Sequence length	5~60	27	27	27	27	27
Neurons	16	2~20	16	16	16	16
Learning rate	0.006	0.006	0.001~0.009	0.001~0.009	0.001~0.009	0.006
Subsequences order	No shuffle	No shuffle	No shuffle	Shuffle	No shuffle	No shuffle
LSTM	1	1	1	1	2	1
Ratio of testing set	15/85	15/85	15/85	15/85	15/85	15/85, 30/85, 45/85

3.3. Rolling prediction method

Inspired by the dynamic forecasting scenario mentioned in (Sagheer and Kotb, 2019), which utilized previous prediction value to calculate the next petroleum production. The proposed deep long short-term memory (DLSTM) model using this method not only had an excellent performance than other reported methods, but also more practical if the label data of testing set was unseen. Therefore, the designed rolling prediction method in this paper can be described that the previous forecasted result is regarded as part of historical data to predict the next mass flow rate. More specifically, the last sequence of historical data in validation set, namely X_t = ($x_{t\text{-sequence\_length}+1}$, $x_{t\text{-sequence\_length}+2}$, $x_{t\text{-sequence\_length}+3}$, ..., x_{t-1}, x_t) as shown in Fig. 4, is sent into the trained model to obtain prediction of time t+1 ($\hat{y}_{t+1}$). Then, the sequence will drop the first data ($x_{t\text{-sequence\_length}+1}$) in X_t and add the predicted value $\hat{y}_{t+1}$ to form a new sequence X_{t+1} ($x_{t\text{-sequence\_length}+2}$, $x_{t\text{-sequence\_length}+3}$, ..., x_{t-1}, x_t, $\hat{y}_{t+1}$) as input to predict flow rate at time t+2 ($\hat{y}_{t+2}$). In this way, the rolling prediction is realized, which is more reasonable than one-step prediction using testing label during forecasting.

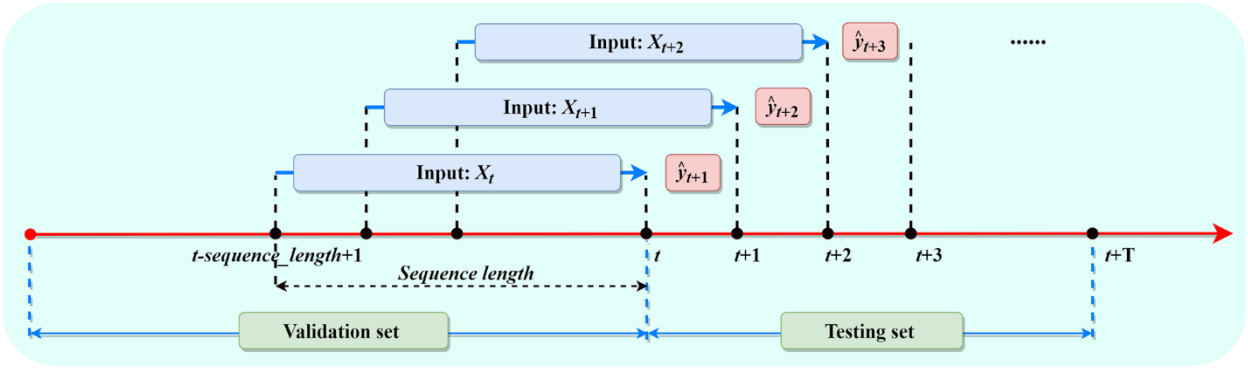

Fig. 4. Schematic plot of rolling prediction.

3.4. Evaluation standards

In order to evaluate the predictive ability of the developed model, root mean square error (RMSE)

and mean absolute percentage error (MAPE) were used as evaluation standards. They are obtained by calculating the overall averaged error between real and predicted mass flow rates in the testing time range. The calculation formulas are shown as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\hat{y}_t - y_t)^2} \tag{9}$$

$$\mathrm{MAPE} = \frac{100\%}{\mathrm{N}}\sum_{t=1}^{N}\left|\frac{\hat{y}_t - y_t}{y_t}\right| \tag{10}$$

Where N is the number of testing set. $\hat{y}_t$ and y_t are the predicted and actual value at time t, respectively.

4. Results and discussions

4.1. Effects of model and training parameters

(1) Sequence length, neurons, learning rate, and shuffle method

Case 1 investigated the effect of sequence length on model prediction ability. As can be seen from Fig. 5a, RMSE and MAPE are generally high when the sequence length is greater than 45. The proposed model trained with sequence length of 27 outperforms due to the lowest values of RMSE and MAPE, which are 1.4668×10^{-4} and 0.2427, respectively. Case 2 trained the developed model with different LSTM neurons. When the number of neurons equals 16, the evaluation standards (i.e., RMSE, MAPE) are the lowest as shown in Fig. 5b. In case 3, the sequence length and number of LSTM neurons were 27 and 16, respectively, and the effect of learning rate on the prediction ability of mass flow rates was studied. It can be seen from Fig. 5c that the rolling prediction ability shows a distinct change with the increase of learning rate, and RMSE and MAPE synchronously reach the minimum (i.e., 1.4969×10^{-4} and 0.2406) when the learning rate is 0.006.

For case 4, the whole sequence array was shuffled by row before dividing into training and validation set. As shown in Fig. 5d, the prediction ability is the best when the learning rate is 0.003, since it has the lowest MAPE and RMSE (1.6907×10^{-4} and 0.2657). Moreover, comparing the best prediction curve of case 3 and case 4, i.e., red line and green line shown in Fig. 6a, it can be found that shuffling subsequences reduces the model's expressiveness. Therefore, the subsequences originating from the training and validation data will not be shuffled in all the other cases. The same evaluation process was also applied to the CFD dataset of 723 K, and a similar conclusion can be obtained as shown in Figs. S3 and 6b.

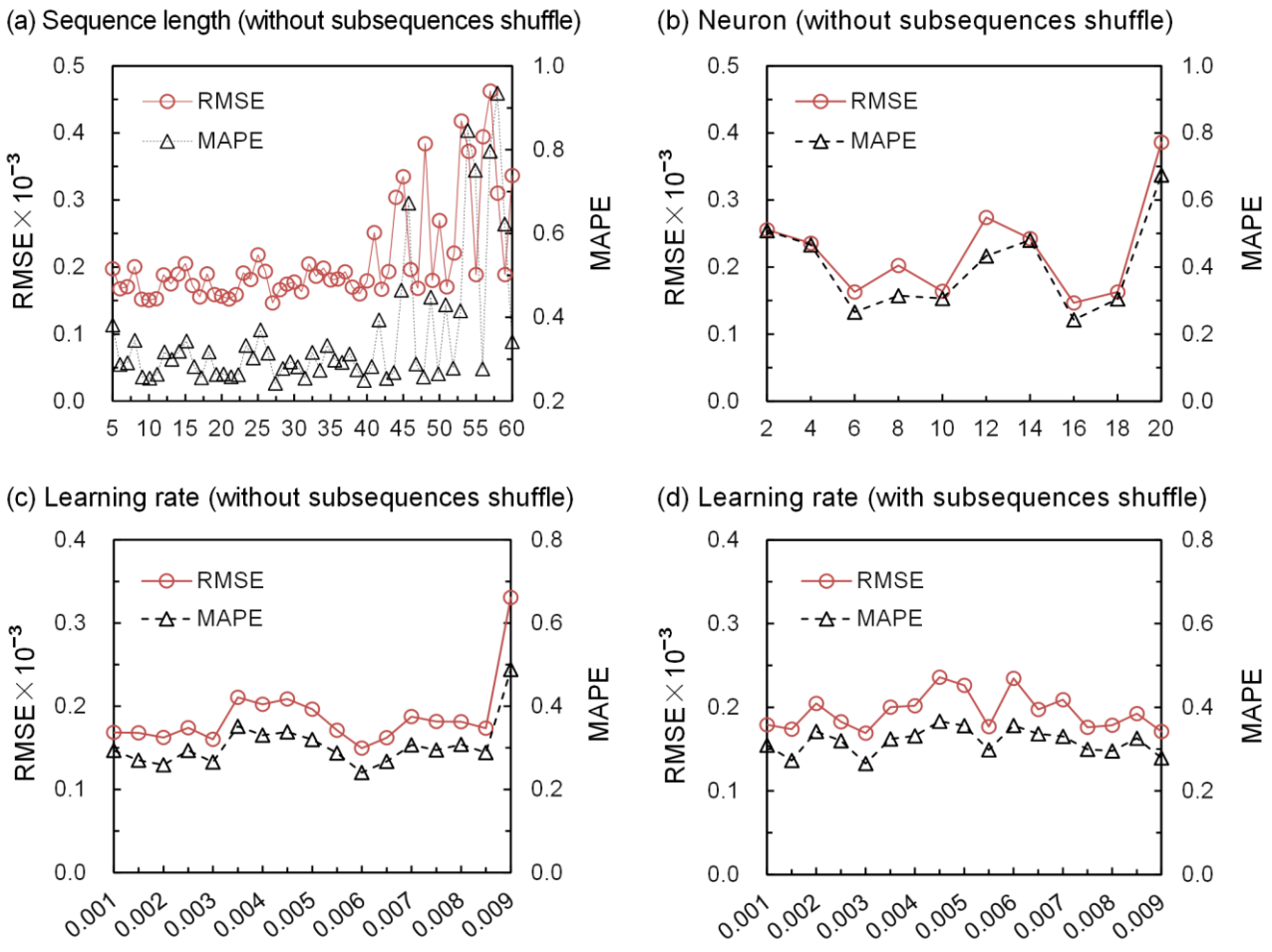

Fig. 5. Effects of sequence length, neurons, learning rate, and shuffle method on RMSE and MAPE at 773 K.

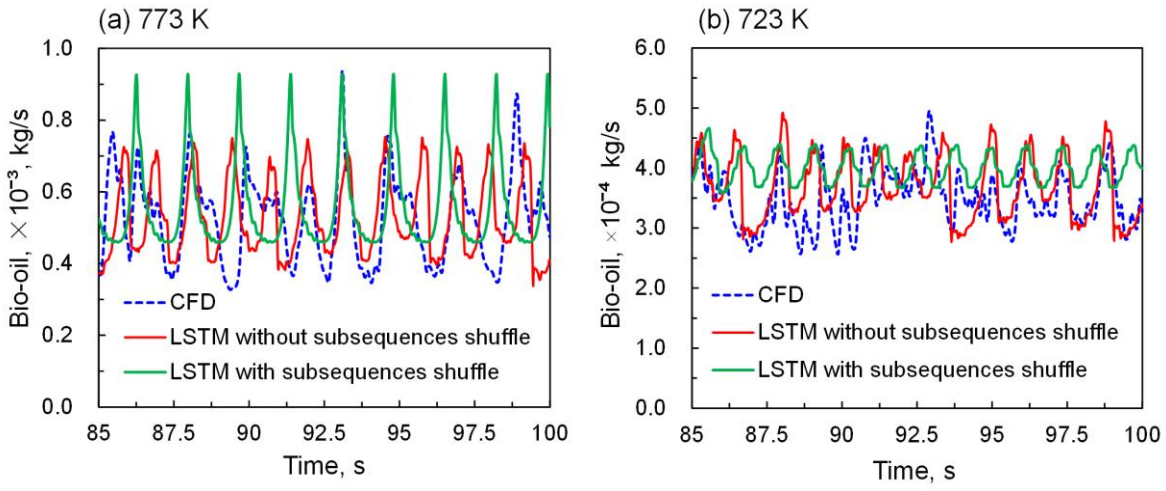

Fig. 6. The predicted mass flow rates of bio-oil at outlet.

(2) Number of LSTM layer

The influence of LSTM layer (i.e., 1 or 2) was investigated in case 5. The developed model framework with two-layer LSTM and its prediction results are shown in Figs. 7 and 8, respectively. In Fig. 8, n1-0.006 curve refers to the predicted mass flow rates by the best single-layer LSTM model with a learning rate of 0.006 in case 3, and n2 curves display the forecasting mass flow rates of double-layer LSTM models trained with different learning rates. It can be found that the predictive ability of single-layer LSTM is far superior to that of double-layer LSTM under the same learning rate (0.006), and the accuracy of double-layer LSTM model is still not good enough even vary the learning rate 0.001 to 0.009. Thus, this double-layer LSTM is not suitable for forecasting mass flow rates in this work. Although more LSTM layers may improve the performance of neural network, the characteristics of sequence data may also be lost due to the increase of the number of layers, which leads to poor performance. The similar phenomenon was also reported by other references. For instance, Wang et al. (2020a) indicated that the model performance of single LSTM layer was much better than that of double LSTM layers for the quality prediction of methyl

methacrylate and vinyl acetate copolymerization process.

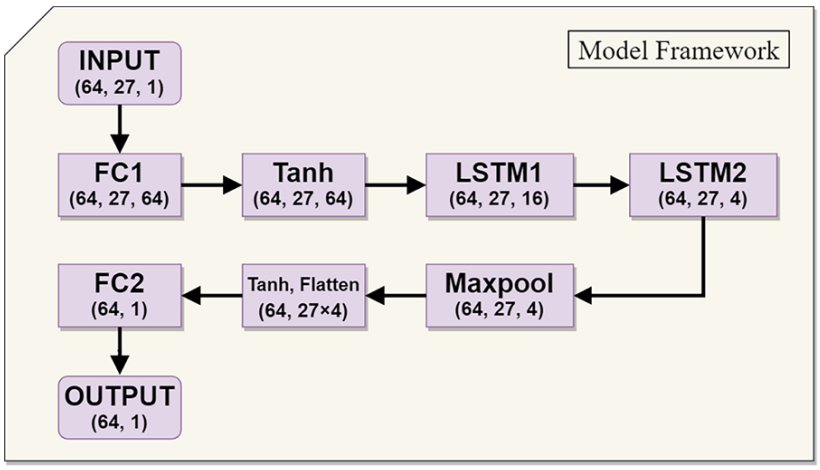

Fig. 7. Model framework of double-layer LSTM.

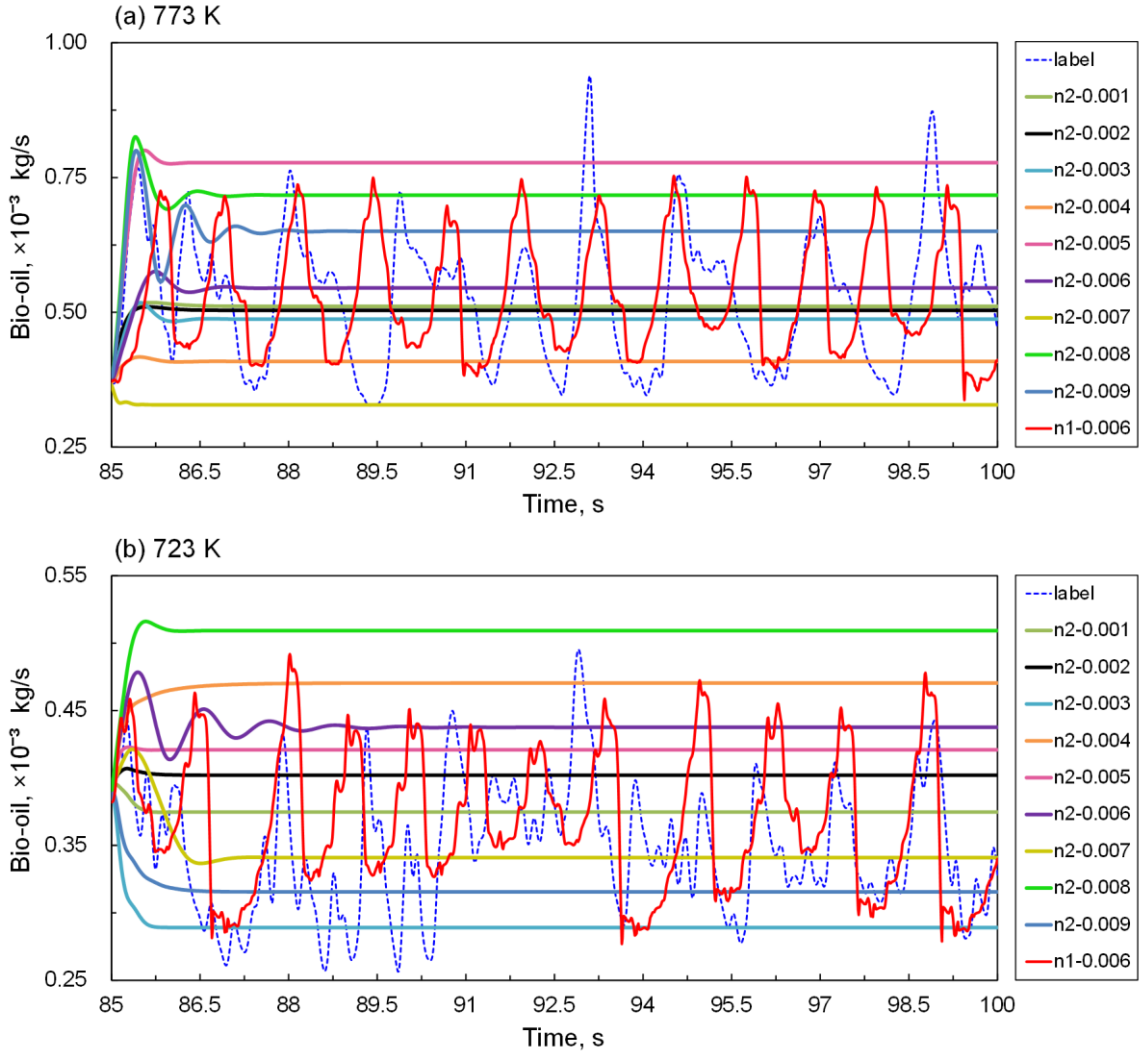

Fig. 8. Effects of LSTM layer and learning rate on the predicted mass flow rates.

(3) Ratio of testing set

Larger ratio of testing set means less CFD computation time for the same investigated time range. Therefore, in case 6, three different ratios were evaluated, i.e., 15/85, 30/85, and 45/85, with the optimal model parameters obtained above. Since the total sample number of mass flow rates are the same, there is less data for model training and validating with the increase of ratio of testing set. The prediction curves are shown in Fig. 9. Obviously, ratios of 15/85 and 30/85 are smaller enough to ensure sufficient number of training and verification data to obtain a LSTM model with good accuracy, as shown in Figs. 9a and 9b. While for the ratio of testing set as large as 45/85, the trained model only expresses a periodic trend of mass flow rates without capturing the detailed vibration (Fig. 9c). Hence, the ratio of testing set is set to 30/85 considering both model accuracy and the ability to reduce CFD simulation time.

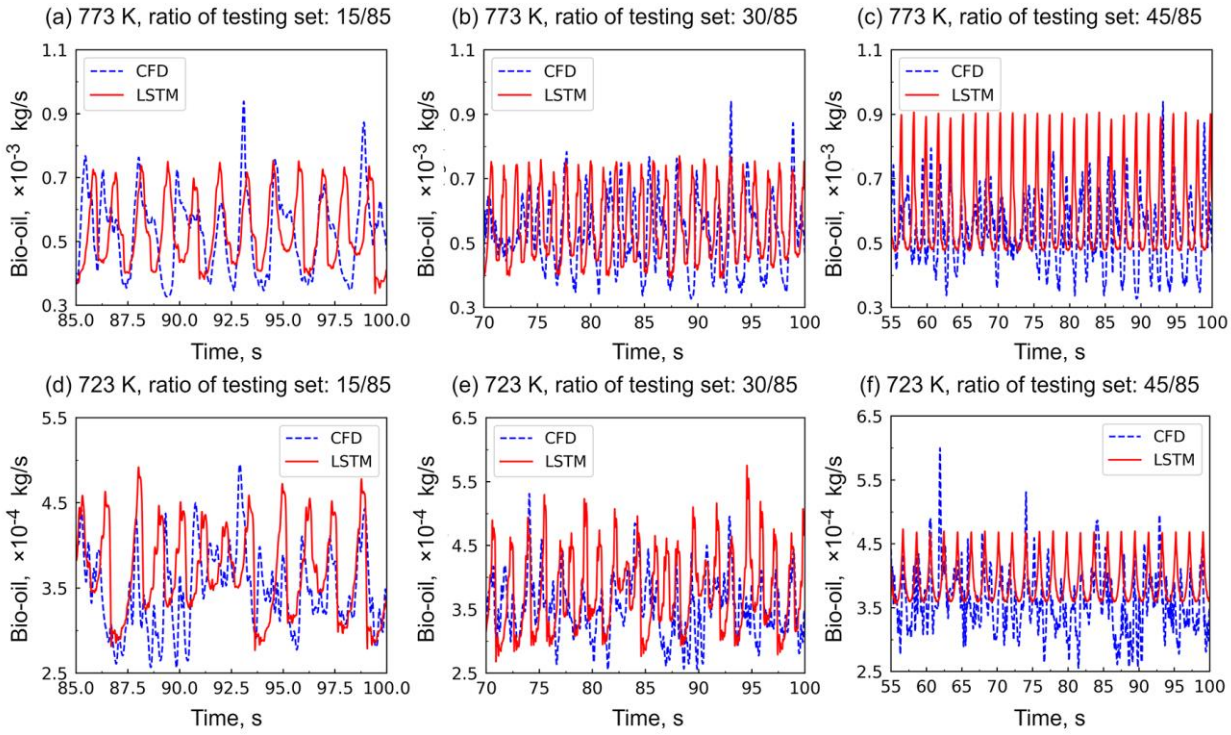

Fig. 9. Influence of ratio of testing set on the predicted mass flow rates.

4.2. Model application to other species

The optimal model configurations obtained in the above section were applied to forecast mass flow rates of other species, i.e., gas, biomass, and char. Firstly, the optimal model was directly used to predict mass flow rates. Then, a fine-tuning process was applied to improve model accuracy, in which the developed model was retrained with new data, and the sequence length and learning rate were optimized, while keeping all the other parameters unchanged. Table 4 shows the value of evaluation standards and relevant parameters before and after fine-tuning, and Fig. 10 compares the CFD results and LSTM predictions before and after fine-tuning.

For all the species at 773 K and gas at 723 K, because their fluctuating characteristics are similar to that of bio-oil, the optimal model before fine-tuning shows reasonable predictions although the magnitude is significantly different. Moreover, the predictive ability can be further improved after fine-tuning as can be seen from the evident decrease of RMSE and MAPE after retraining as shown in Table 4. In this situation, the sequence length and learning rate keep constant before and after fine-tuning, which means the optimal parameters can be used for the dataset with similar fluctuating characteristics. Only retraining with new dataset is needed.

However, because the fluctuating characteristics of biomass and char at 723 K are drastically different with that of bio-oil, the model before fine-tuning can hardly capture the trends of vibration. After fine-tuning, the RMSE and MAPE decrease, and the fluctuating characteristics are reproduced much better. The learning rate is still 0.006 for biomass, while it increases slightly to 0.0085 for char. Nevertheless, the sequence length increases from 27 to 50 for both biomass and char. This is because that LSTM only learns the fluctuating characteristics in each sequence.

Consequently, the sequence length should be larger for datasets with lower-frequency fluctuations (biomass and char at 723 K).

Table 4 Parameters and evaluation standards before and after fine-tuning.

T, K	Species	Before fine-tuning				After fine-tuning			
		Sequence length	Learning rate	RMSE	MAPE	Sequence length	Learning rate	RMSE	MAPE
773	Gas	27	0.006	3.083×10^{-5}	0.2656	27	0.006	2.982×10^{-5}	0.2414
	Biomass	27	0.006	3.638×10^{-5}	0.5224	27	0.006	3.532×10^{-5}	0.4753
	Char	27	0.006	3.105×10^{-5}	0.5366	27	0.006	2.648×10^{-5}	0.4196
723	Gas	27	0.006	1.566×10^{-5}	0.2063	27	0.006	1.296×10^{-5}	0.1646
	Biomass	27	0.006	1.945×10^{-4}	0.5911	50	0.006	1.381×10^{-4}	0.4226
	Char	27	0.006	3.273×10^{-5}	0.5677	50	0.0085	2.994×10^{-5}	0.4227

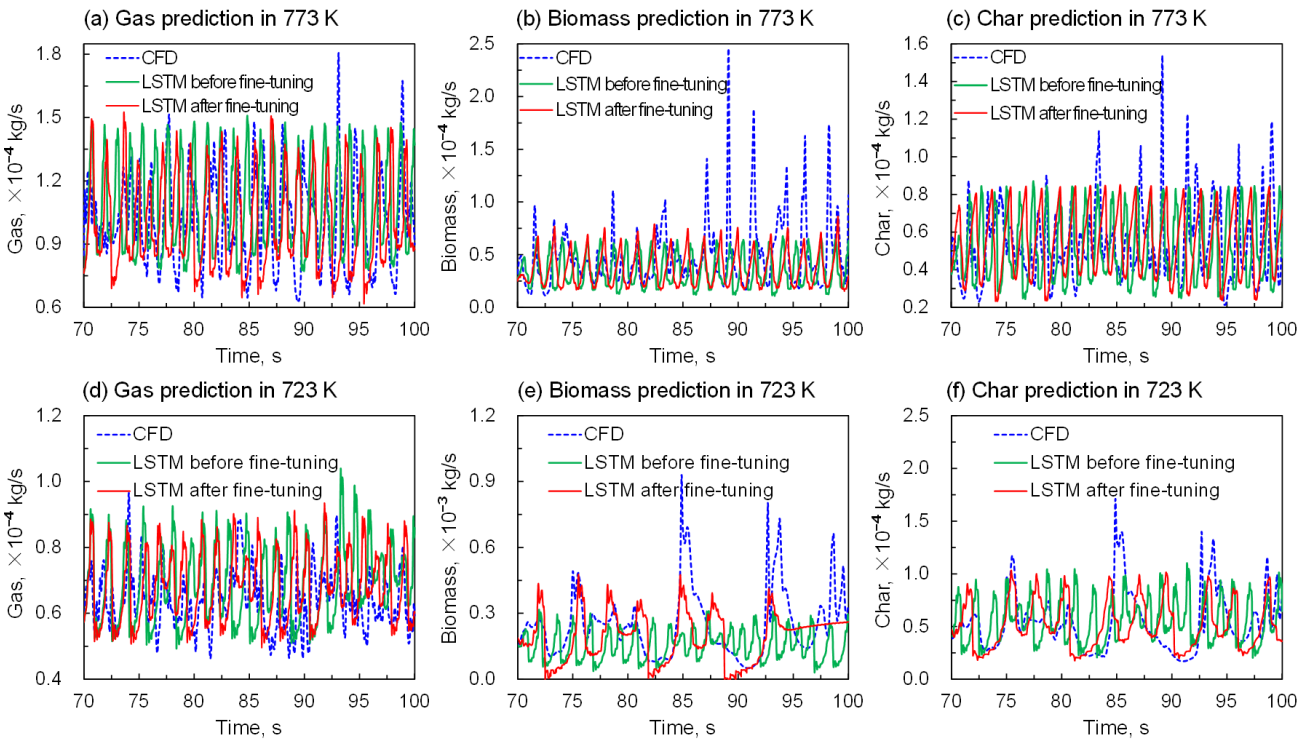

Fig. 10. Predicted results of gas, biomass, and char before and after fine-tuning.

4.3. Assessment of product yields

The product yields predicted by LSTM are evaluated by integrating the mass flow rates over the last 30 s in the CFD simulation as shown in Figs. 11 and 12. Since the flow rate of feeding biomass particles is constant, Table 5 only gives the integrated mass of all species over the last 30 s. Apparently, the developed LSTM performs fairly well on bio-oil, gas, and char, with low relative errors of 4.83%, -1.61%, and 2.27% at 773 K, while they are 6.69%, 5.80%, and 4.89% at 723 K, respectively, although the predictions are all relatively poor for biomass at both temperatures. In this work, only LSTM network was used to forecast mass flowrates by a rolling prediction method. In our future work, advanced methods and/or other neural networks such as attention mechanism and convolutional neural network (CNN) will be coupled into the current LSTM network to improve the accuracy of each data point, especially for biomass. Through the proposed LSTM method, the CFD simulation time is reduced by nearly 30%, which will be especially practical for optimizing operating conditions or larger-scale simulation systems. Even for this work with laboratory-scale reactor, the training process typically costs 1 min 30 s, and it takes only 40 s to predict flow rates in the future 30 s. However, 18 h is needed for the CFD simulation of the last 30 s, which confirms that the proposed method in this work is able to reduce lots of computational effort. In addition, the well-captured fluctuating characteristics of mass flow rates and predicted final product yields have great potential to improve the accuracy of process simulation for whole unit or plant, which can facilitate the development of digital twin and smart factories.

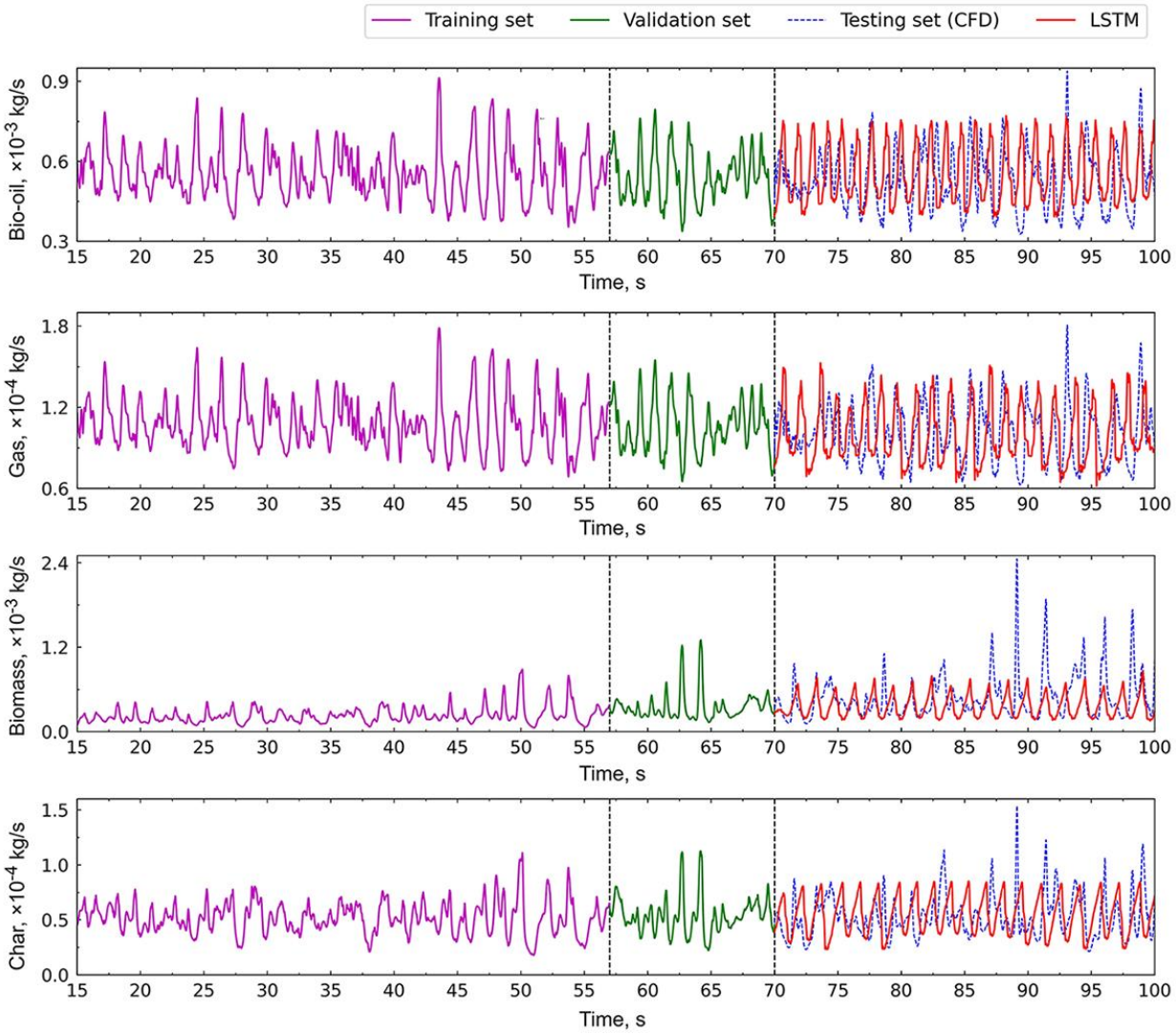

Fig. 11. The mass flow rates of CFD and LSTM for all species at 773 K.

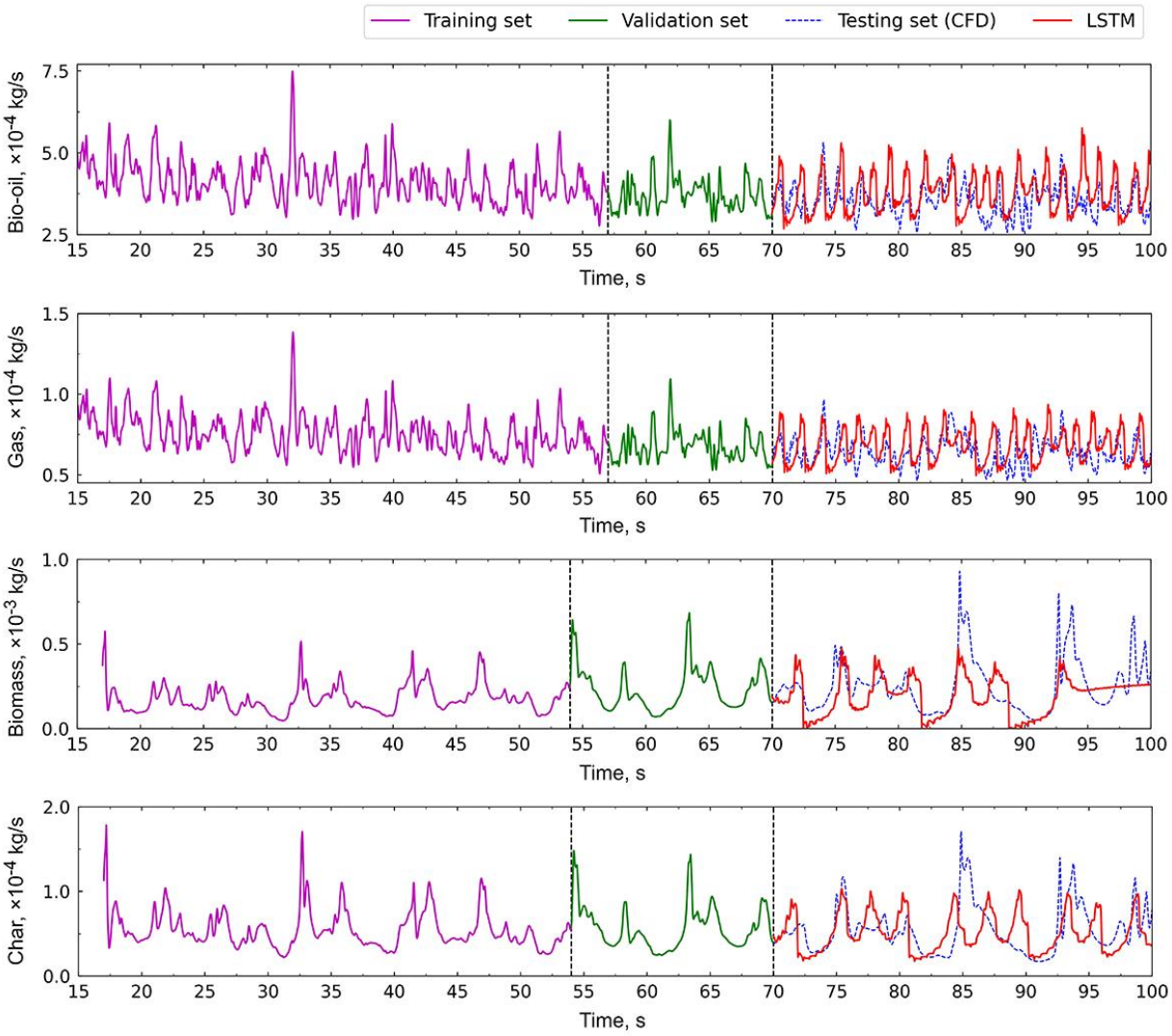

Fig. 12. The mass flow rates of CFD and LSTM for all species at 723 K.

Table 5 The integrated mass of CFD and LSTM over the last 30 s.

T, K	Species	CFD predicted mass, $\times 10^{-3}$ kg	LSTM predicted mass, $\times 10^{-3}$ kg	Relative error, %
773	Bio-oil	15.8216	16.5859	4.83
	Gas	3.0422	2.9932	-1.61
	Biomass	1.4491	1.0935	-24.54
	Char	1.5958	1.6320	2.27
723	Bio-oil	10.5961	11.3051	6.69
	Gas	1.9208	2.0322	5.80
	Biomass	7.5106	5.8371	-22.28
	Char	1.6234	1.5440	-4.89

5. Conclusions

In this paper, the LSTM method was used to predict the mass flow rates at reactor outlet for biomass fast pyrolysis in a bubbling fluidized bed through training the historical CFD data from MFM simulation. Firstly, the optimal parameters were determined by investigating six main influencing factors on the dataset of bio-oil, i.e., sequence length, number of neurons, learning rate, subsequences order (shuffle or not), number of LSTM layers, and ratio of testing set, which is 27, 16, 0.006, no shuffle, 1, and 30/85, respectively. Then, the mass flow rates of other species, including gas, biomass, and char were predicted using the developed LSTM model. In addition, CFD data from a different pyrolysis temperature (723 K) was also tested. For the dataset with similar fluctuating characteristics, good predictions can be obtained directly, and retraining with new dataset is recommended to achieve better performance. However, the sequence length should

be larger for dataset of biomass and char at 723 K with lower-frequency fluctuations. Overall, the predicted product yields of most species were in good agreement with those originated from CFD simulation, while reducing nearly 30% CFD simulation time. Moreover, the well-predicted fluctuating characteristics and final product yields are valuable for the process simulation of a whole unit or plant, which is helpful for digitalizing key reactors and building smart factories.

Acknowledgements

Financial supports from the Natural Science Basic Research Program of Shaanxi (Program No. 2023-JC-YB-119), the National Natural Science Foundation of China (No. 22178123), and the Postgraduate Innovation and Practice Ability Development Fund of Xi'an Shiyou University (YCS21211039) were greatly appreciated.

References

Arora, P., Hoadley, A.F.A., Mahajani, S.M., Ganesh, A., 2017. Compartment model for a dual fluidized bed biomass gasifier. Chem. Eng. Res. Des. 117, 274-286. http://doi.org/10.1016/j.cherd.2016.10.025.

Aversano, G., Ferrarotti, M., Parente, A., 2021. Digital twin of a combustion furnace operating in flameless conditions: reduced-order model development from CFD simulations. P. Combust. Inst. 38(4), 5373-5381. http://doi.org/10.1016/j.proci.2020.06.045.

Bazai, H., Kargar, E., Mehrabi, M., 2021. Using an encoder-decoder convolutional neural network to predict the solid holdup patterns in a pseudo-2d fluidized bed. Chem. Eng. Sci. 246, 116886. http://doi.org/10.1016/j.ces.2021.116886.

Ding, J., Gidaspow, D., 1990. A bubbling fluidization model using kinetic theory of granular flow. AICHE J. 36(4), 523-538. http://doi.org/10.1002/aic.690360404.

Fakayode, O.A., Wahia, H., Zhang, L., Zhou, C., Ma, H., 2023. State-of-the-art co-pyrolysis of lignocellulosic and macroalgae biomass feedstocks for improved bio-oil production- A review. Fuel 332, 126071. http://doi.org/10.1016/j.fuel.2022.126071.

Houston, R., Oyedeji, O., Abdoulmoumine, N., 2022. Detailed biomass fast pyrolysis kinetics integrated to computational fluid dynamic (CFD) and discrete element modeling framework: Predicting product yields at the bench-scale. Chem. Eng. J. 444, 136419. http://doi.org/10.1016/j.cej.2022.136419.

Huilin, L., Gidaspow, D., 2003. Hydrodynamics of binary fluidization in a riser: CFD simulation using two granular temperatures. Chem. Eng. Sci. 58(16), 3777-3792. http://doi.org/10.1016/S0009-2509(03)00238-0.

Kim, T.-H., Choi, M.K., Choi, H.S., 2022. Biomass fast pyrolysis prediction model through data-based prediction models coupling with CPFD simulation. J. Anal. Appl. Pyrolysis 162, 105448. http://doi.org/10.1016/j.jaap.2022.105448.

Kochkov, D., Smith, J.A., Alieva, A., Wang, Q., Brenner, M.P., Hoyer, S., 2021. Machine learning-accelerated computational fluid dynamics. Proc. Natl. Acad. Sci. USA 118(21), e2101784118. http://doi.org/10.1073/pnas.2101784118.

Ladický, L.u., Jeong, S., Solenthaler, B., Pollefeys, M., Gross, M., 2015. Data-driven fluid simulations using regression forests. ACM. T. Graphic. 34(6), Article 199. http://doi.org/10.1145/2816795.2818129.

Lao, Z., Shao, Y., Gao, X., 2022. Multiscale CFD Modeling of High-Temperature Biomass Pyrolysis with an Intraparticle Particle Model and Detailed Pyrolysis Kinetics. Ind. Eng. Chem. Res. 61(45), 16843-16856. http://doi.org/10.1021/acs.iecr.2c02992.

Lu, L., Gao, X., Dietiker, J.-F., Shahnam, M., Rogers, W.A., 2021. Machine learning accelerated discrete element modeling of granular flows. Chem. Eng. Sci. 245, 116832. http://doi.org/10.1016/j.ces.2021.116832.

Lu, L., Gao, X., Dietiker, J.-F., Shahnam, M., Rogers, W.A., 2022. MFiX based multi-scale CFD simulations of biomass fast pyrolysis: A review. Chem. Eng. Sci. 248, 117131. http://doi.org/10.1016/j.ces.2021.117131.

Ouyang, B., Zhu, L., Luo, Z., 2022a. Machine learning for full spatiotemporal acceleration of gas-particle flow simulations. Powder Technol. 408, 117701. http://doi.org/10.1016/j.powtec.2022.117701.

Ouyang, Y., Vandewalle, L.A., Chen, L., Plehiers, P.P., Dobbelaere, M.R., Heynderickx, G.J., Marin, G.B., Van Geem, K.M., 2022b. Speeding up turbulent reactive flow simulation via a deep artificial neural network: A methodology study. Chem. Eng. J. 429, 132442. http://doi.org/10.1016/j.cej.2021.132442.

Porrazzo, R., White, G., Ocone, R., 2016. Fuel reactor modelling for chemical looping combustion: From micro-scale to macro-scale. Fuel 175, 87-98. http://doi.org/10.1016/j.fuel.2016.01.041.

Sagheer, A., Kotb, M., 2019. Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing 323, 203-213. http://doi.org/10.1016/j.neucom.2018.09.082.

Shoaib Ahmed Khan, M., Grioui, N., Halouani, K., Benelmir, R., 2022. Techno-economic analysis of production of bio-oil from catalytic pyrolysis of olive mill wastewater sludge with two different cooling mechanisms. Energy. Convers. Man-X. 13, 100170. http://doi.org/10.1016/j.ecmx.2021.100170.

Silvestri, L., 2021. CFD modeling in Industry 4.0: New perspectives for smart factories. Procedia Comput. Sci. 180, 381-387. http://doi.org/10.1016/j.procs.2021.01.359.

Sun, W., Huang, C., 2020. A carbon price prediction model based on secondary decomposition algorithm and optimized back propagation neural network. J. Clean. Prod. 243, 118671. http://doi.org/10.1016/j.jclepro.2019.118671.

Syamlal, M., 1987. The Particle-Particle Drag Term in a Multiparticle Model of Fluidization, National Technical Information Service, Springfield, VA. DOE/MC/21353-2373, NTIS/DE87006500.

Vikram, S., Rosha, P., Kumar, S., 2021. Recent Modeling Approaches to Biomass Pyrolysis: A Review. Energy Fuel 35(9), 7406-7433. http://doi.org/10.1021/acs.energyfuels.1c00251.

Wang, K., Gopaluni, R.B., Chen, J., Song, Z., 2020a. Deep Learning of Complex Batch Process Data and Its Application on Quality Prediction. IEEE T. Ind. Inform. 16(12), 7233-7242. http://doi.org/10.1109/tii.2018.2880968.

Wang, Q., Guo, Y., Yu, L., Li, P., 2020b. Earthquake Prediction Based on Spatio-Temporal Data Mining: An LSTM Network Approach. IEEE T Emerg. Top. Com. 8(1), 148-158. http://doi.org/10.1109/tetc.2017.2699169.

Wu, F., Li, F., Zhao, X., Bolan, N.S., Fu, P., Lam, S.S., Mašek, O., Ong, H.C., Pan, B., Qiu, X.,

Rinklebe, J., Tsang, D.C.W., Van Zwieten, L., Vithanage, M., Wang, S., Xing, B., Zhang, G., Wang, H., 2022. Meet the challenges in the "Carbon Age". Carbon Res. 1(1), 1. http://doi.org/10.1007/s44246-022-00001-9.

Xie, Z., Gu, X., Shen, Y., 2022. A Machine Learning Study of Predicting Mixing and Segregation Behaviors in a Bidisperse Solid–Liquid Fluidized Bed. Ind. Eng. Chem. Res. 61(24), 8551-8565. http://doi.org/10.1021/acs.iecr.2c00071.

Xiong, Q., Yang, Y., Xu, F., Pan, Y., Zhang, J., Hong, K., Lorenzini, G., Wang, S., 2017. Overview of Computational Fluid Dynamics Simulation of Reactor-Scale Biomass Pyrolysis. ACS. Sustain. Chem. Eng. 5(4), 2783-2798. http://doi.org/10.1021/acssuschemeng.6b02634.

Xue, Q., Dalluge, D., Heindel, T., Fox, R., Brown, R., 2012. Experimental validation and CFD modeling study of biomass fast pyrolysis in fluidized-bed reactors. Fuel 97, 757-769. http://doi.org/10.1016/j.fuel.2012.02.065.

Zhao, J., Deng, F., Cai, Y., Chen, J., 2019. Long short-term memory - Fully connected (LSTM-FC) neural network for PM2.5 concentration prediction. Chemosphere 220, 486-492. http://doi.org/10.1016/j.chemosphere.2018.12.128.

Zhong, H., Liang, S., Zhang, J., Zhu, Y., 2016a. Multi-fluid model with variable particle density and diameter based on mass conservation at the particle scale. Powder Technol. 294, 43-54. http://doi.org/10.1016/j.powtec.2016.02.024.

Zhong, H., Xiong, Q., Yin, L., Zhang, J., Zhu, Y., Liang, S., Niu, B., Zhang, X., 2020. CFD-based reduced-order modeling of fluidized-bed biomass fast pyrolysis using artificial neural network. Renew. Energy 152, 613-626. http://doi.org/10.1016/j.renene.2020.01.057.

Zhong, H., Xiong, Q., Zhu, Y., Liang, S., Zhang, J., Niu, B., Zhang, X., 2019. CFD modeling of the effects of particle shrinkage and intra-particle heat conduction on biomass fast pyrolysis. Renew. Energy 141, 236-245. http://doi.org/10.1016/j.renene.2019.04.006.

Zhong, H., Zhang, J., Zhu, Y., Liang, S., 2016b. Multi-fluid modeling biomass fast pyrolysis in the fluidized-bed reactor including particle shrinkage effects. Energy Fuel 30(8), 6440-6447. http://doi.org/10.1021/acs.energyfuels.6b00914.

Zhu, L., Chen, X., Ouyang, B., Yan, W., Lei, H., Chen, Z., Luo, Z., 2022. Review of Machine Learning for Hydrodynamics, Transport, and Reactions in Multiphase Flows and Reactors. Ind. Eng. Chem. Res. 61(28), 9901-9949. http://doi.org/10.1021/acs.iecr.2c01036.

Zhu, L., Tang, J., Luo, Z., 2020. Machine learning to assist filtered two-fluid model development for dense gas–particle flows. AICHE J. 66(6), e16973. http://doi.org/10.1002/aic.16973.