Journal of Physics:
Conference Series

PAPER • **OPEN ACCESS**

Fluidized Bed Reactors for energy recovery from biomass and wastes: a data-driven approach towards the development of digital twins for real time control and monitoring

To cite this article: Matteo Baldelli *et al* 2024 *J. Phys.: Conf. Ser.* **2893** 012014

Journal of Physics: Conference Series **2893** (2024) 012014 doi:10.1088/1742-6596/2893/1/012014

Fluidized Bed Reactors for energy recovery from biomass and wastes: a data-driven approach towards the development of digital twins for real time control and monitoring

Matteo Baldelli[1], **Lorenzo Bartolucci**[1], **Stefano Cordiner**[1] **and Vincenzo Mulone**[1]

[1]Department of Industrial Engineering, Tor Vergata University of Rome, via del Politecnico 1, 00133, Rome, Italy

E-mail: matteo.baldelli@uniroma2.it

Abstract. The application of machine learning (ML) techniques for the control and development of digital twins for a fluidized bed reactors represents a significant advancement in process engineering. In this study, the integration of data-driven models trained using computational fluid dynamics (CFD) simulations, is explored for developing and optimizing the lab-scale fluidized bed reactor operations. By leveraging the collection of data generated from CFD simulations, data-driven algorithms, based on the Singular Value Decomposition (SVD) and Gaussian processes for regression, are trained to predict the gas-solid flow patterns under different operating condition. The data-driven models presented, serve as efficient reduced order model (ROM) surrogate for computationally expensive CFD simulations, enabling real-time predictions and control strategies for fluidized bed reactors, facilitating continuous monitoring, optimization, and predictive maintenance.Moreover, the ROM can effectively capture the complex relationships within the reactor system, with an overall error $< 10\%$ even without precise knowledge of the underlying physical phenomena. The synergistic combination of ML techniques and CFD simulations offers valuable insights into complex multiphase flow phenomena and reactor dynamics, leading to improved process control, energy efficiency, and overall performance of fluidized bed reactors. This approach holds great promise for accelerating innovation and sustainability in chemical and energy industries.

1 Introduction

The recovery of energy, fuels, and chemicals from residual biomass and plastic waste plays a crucial role in environmental sustainability and resource scarcity challenge [1, 2]. Pyrolysis is one of the most studied technologies to transform waste into valuable resources [3, 4], due to its potential products flexibility. Despite this potential capability, these system are often rigid and designed for specific waste types [5]. However, flexible pyrolysis system, in terms of the feedstock and desired output, could solve several problems related to the waste recycling. In particular, they could be adapted to handle a wider range of feedstock, and even seasonal fluctuation in waste availability, maximizing their efficiency and resource recovery potential [6, 7, 8]. This adaptability allows also to process different waste streams that might otherwise be difficult or expensive to manage, increasing the economic viability of this

waste-to-energy processes. From the point of view of the final output, pyrolysis reactors can be tailored to produce a specific output depending on market demands and application requirements. To achieve such flexibility, advanced real-time control devices are essential. These devices must posses a comprehensive knowledge of the reactors' response for different feedstock compositions and for different operating conditions, and actuate, in real-time application, the appropriate actions at the right time. In such engineering applications, accurate physical systems modeling is mandatory for designing reliable control strategies. Complex physical system like pyrolysis reactors, can be effectively described only by high-fidelity simulations (HFSs). HFSs are usually expensive and computationally demanding and, due to the non linearities of these problems, a slightly changing in the operating condition can drastically modified the state of the system. Therefore, a complete knowledge about the system's behavior can be reached by performing an intense numerical simulation campaign for all the possible combination of the inputs parameters, until the desired number of observation is obtained.
In this work, we proposed a reduced order model (ROM) base on data, of a lab-scale fluidized bed reactor for the pyrolysis, focusing on the fluid-mechanics point of view. In particular, our goal is to propose a surrogate model able to accurately predict the behavior of the reactor, without the need of performing computational fluid dynamics (CFD) simulations [9]. ROMs are a reliable way to develop digital twins of real systems, enabling real-time control and monitoring applications [10, 11, 12]. Usually, ROM are mathematical model based on data that try to understand and approximate the underlying relationship between input parameters and the resulting physical phenomena. Several approach can be found in literature related to Machine Learning (ML) techniques applied in fluid dynamics and in reacting flow prediction and control [13, 9, 14, 15]. In pyrolysis application, data-driven approach are commonly used to predict the final product the reactor or in general to infer the kinetics of chemical reactions [16, 17, 18, 19, 20]. In those applications, ROMs are generally developed using techniques that directly connect the input parameters (*e.g.* feedstock flow rate, pyrolysis temperature) to the output. These approaches usually required a very large amount of data since no prior knowledge is considered and, moreover, the relationship between input and output might be lost during the training. The number of training data could be reduced introducing some physical constraints based on a some prior knowledge about the physics phenomena (physics informed approach) [21, 22, 23]. In some cases, is possible that the original data-set could be reduced and represented by a fewer set of variables reducing also the number of data needed and ROMs to train. This principle summarized the basic idea of this work.
Here, we proposed the development of a ROM for fluidized-bed reactor through fluid dynamics phenomena occurring within the reactor, based on problem dimensionality reduction [24, 25, 26, 27]. Singular Value Decomposition (SVD) is a mathematical techniques for decomposing matrices used to find a set of orthogonal and low dimensional basis functions able to represent high-dimensional data. These basis functions are used to project the original data matrix in a low-dimensional space (compression), finding a new set of uncorrelated variable that are able to completely described the original data set. In this study, the SVD approach is used to reduce the dimensionality of problem, capturing the underlying relationship of the variable, and using the reduced representation of the data together with Gaussian Processes (GP) to infer the behaviour of the fluidized bed reactor for unexplored input parameters. [28, 24]. In particular, we focused on the prediction of three important fluid field: the pressure drop inside the reactor, the velocity of the fluidising gas, and the volume fraction of the bed material.
This paper is organized as follows: in the Section 2 the Methods are described in details; Section 3 presents the Results and Discussions; finally, in Section 4 the Conclusions are illustrated.

2 Methods

The aim of the study here proposed, is to provide a possible approach to develop a ROM of a fluidized-bed reactors using ML approaches. To accomplish this objective, the first step was collecting data about the pressure p, fluidising gas velocity $\mathbf{u}_{N_2}$ and solid volume fraction fields v, from CFD simulations varying the operating condition, namely the inlet velocity of the fluidising gas $\mathbf{u}_{in,N_2}$, of the fluidized-bed system; the second phase was to develop a ROM of the fluid mechanics phenomena that occur within the reactor, based on SVD and dimensionality reduction techniques. In this section, the numerical simulation set-up, SVD method, and the data-driven approach for predicting quantities, are explained.

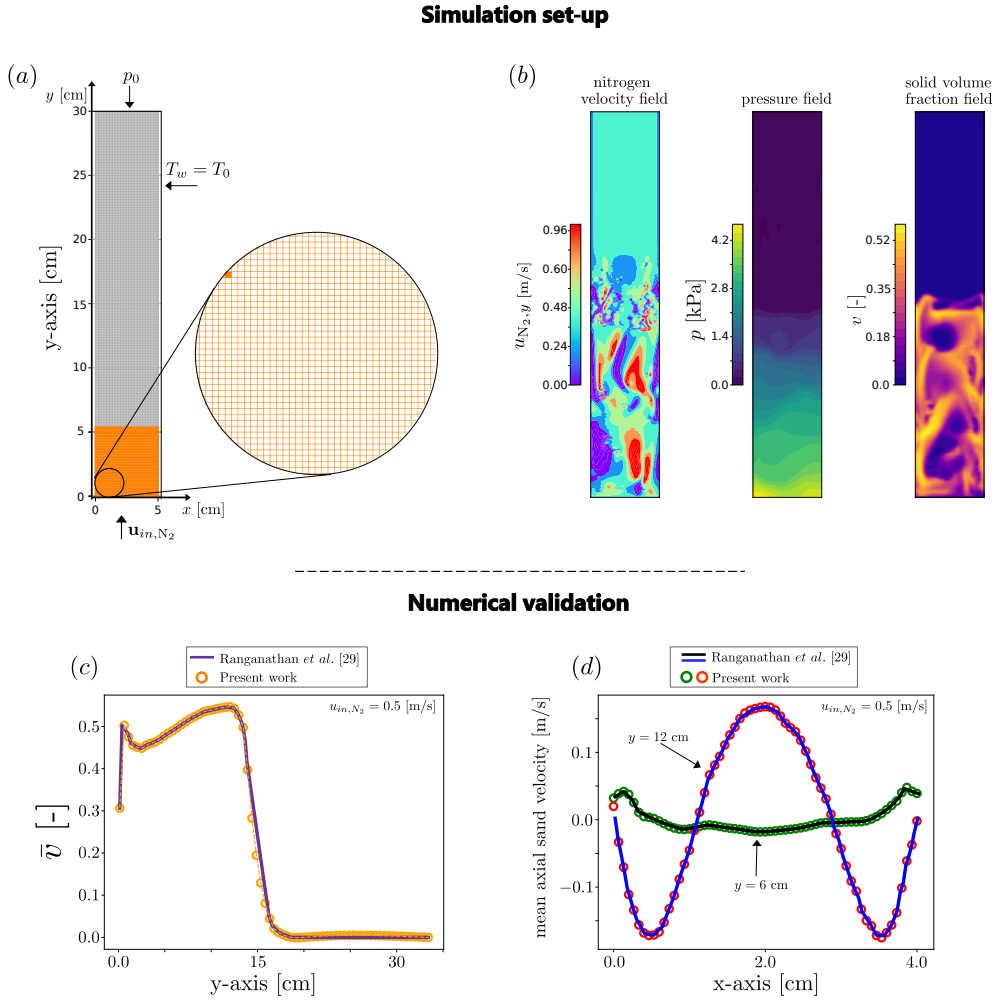

Figure 1: **Simulation set-up and numerical validation of a fluidized-bed reactor.** **(a)** The fluidized-bed reactor is schematized as a 2D rectangular fluid domain with 5 cm width and 30 cm heigh. The fluidising gas is nitrogen and enters in the system with a constant velocity $\mathbf{u}_{in,N_2}$. The reactor top section is a pressure outlet at a fixed pressure $p_0 = 1$ bar. Reactor walls, nitrogen and sand are a constant temperature $T_0 = 500$ °C. The orange area, represent the volume occupied by the bed material at $t_0 = 0$. The inset shows a zoom to highlight the geometry mesh. **(b)** Examples of physical field of interest in this work. From left to right: nitrogen velocity field; pressure field; solid volume fraction field. Each field is presented with a different colormap to allow for better distinction. **(c,d)** Numerical validation of the CFD set-up. The numerical setting used in the present work is applied on a slightly different system and compared with the work of Ranganathan *et al.* [29]. In panel (c), the mean of solid volume fraction along the height of the reactor is reported. In panel (d), the mean axial sand velocity evaluated for two different reactor heights ($y = 6 - 12$ cm) is shown. The solid lines represent the results reported in [29], while the empty circle are the results obtained from the validation of the numerical setting used in the present work.

2.1 Fluidized-bed reactor numerical simulations

In the numerical simulations, the lab-scale fluidized-bed reactor is modeled as a 2D rectangular fluid domain of width 5 cm and height of 30 cm. The domain is then discretized with an uniform mesh, where each element is $5 \times 10^{-}4$ m side square for a total of $m = 60 \times 10^3$ grid points, see Fig. 1(a). In the reactor bottom section (velocity inlet), the fluidising gas, nitrogen (N_2), flows with a constant velocity $\mathbf{u}_{in,N_2}$ and, at the time $t_0 = 0$, encounters 5 cm^2 volume (in 2D sense) of sand (inert bed material) with a density of $\rho_S = 2700$ kg/m^3 and a granular diameter $d_S = 5 \times 10^{-4}$ m. The volume occupied by the sand at the beginning of the simulation is depicted as an orange area in Fig. 1(a). The reactor top section is defined as a pressure outlet, hence a constant pressure is imposed for the entire duration of the simulation. The reactor walls, the nitrogen inflow and the sand are at a constant temperature of $T_0 = 500$ °C. Under this operating conditions the system is allowed to evolve for 27 seconds, *i.e.* the simulation time, and each 0.001s a simulation snapshot is saved. In particular, during each time-step the measures of the pressure drop p within the reactor, the N_2 velocity field $\mathbf{u}_{N_2}$ and solid volume fraction v are saved, see Fig. 1(b). A simulation campaign is then conducted varying the inlet nitrogen velocity $\mathbf{u}_{in,N_2}$ from 0.06 m/s to 0.35 m/s.

From a numerical point of view, in the fluidised-bed reactor simulations here proposed, an Eulerian-Eulerian model for the two phase flow is solved. In particular, one solid phase representing the bed-material and one gaseous phase that models the fluidising gas. Since in this preliminary study the chemical reaction are not taken into account, only the continuity, momentum and the energy equations are solved for each phase separately. The coupling between the two phases, is achieved through the exchanged forces. For the momentum transfer, the Symlal *et al.* drag coefficient is used [30]. For the interphase heat transfer, the Gunn heat transfer coefficient between the solid and gaseous phase is adopted [31]. Since the sand motion is also modeled through granular flow kinetics, the granular bulk viscosity and the granular viscosity of sand are modeled using the Lun *et al.* and Symlal relationship, respectively [32, 33]. The CFD simulations are performed using the commercial software ANSYS® Fluent [34, 35].

Finally, a numerical validation of the CFD simulation set-up used in the present work is summarized in Fig. 1(c,d). The validation of the numerical setting is performed on a slightly different fluidized-bed system (solved using the same numerical set-up here suggested) proposed by Ranganathan *et al.* [29]. The validation results are presented by comparing the mean of the solid volume fraction $\bar{v}$ along the reactor axis and the mean axial velocity of the bed evaluated at two different heights ($y = 6 - 12$ cm). As observed, the proposed numerical model is in excellent agreement with the results shown in [29].

2.2 Singular value decomposition and dimensionality reduction

The strategy used in the present study to build a ROM of a fluidized-bed reactor, is depicted in Fig. 2. Let's consider a high-fidelity simulation (HFS, sometimes called Full Order Model) as described in section 2.1 and suppose to collect the data (spatially and temporally) related to some fields (*e.g* pressure field, velocity field, ...) in a matrix $\mathbf{D} \in \mathbb{R}^{m \times s}$ in which each column $\mathbf{d}_k \in \mathbb{R}^m$, with $k \in [1, ..., s]$, represent those fields evaluated in all the mesh grid points at certain time (snapshot). A complementary interpretation of the data matrix can be obtained by observing the rows rather than the columns. In fact, each row represents the measure of a particular quantity of interest (field) at a fixed point in space as time varies. In essence, given a physical system simulated with a uniform temporal discretization $[t_k = k\Delta t]_{k=0}^{s}$, given the number m of Cartesian grid points $\mathbf{x}_i$ where the field is evaluated, and after reshaping each snapshot in a column vector, the data matrix $\mathbf{D}$ can be written as follow:

$$\mathbf{D} = \begin{bmatrix} \mathbf{d}_1(\mathbf{x}_0, t_0) & \dots & \mathbf{d}_k(\mathbf{x}_0, t_k) & \dots & \mathbf{d}_s(\mathbf{x}_0, t_s) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{d}_1(\mathbf{x}_m, t_0) & \dots & \mathbf{d}_k(\mathbf{x}_m, t_k) & \dots & \mathbf{d}_s(\mathbf{x}_m, t_s) \end{bmatrix},$$

where the column $\mathbf{d}_k \in \mathbb{R}^m$, represents the field measurements in the reshaped mesh. It is important to highlight that the method used to reshape each snapshot into a vector is no relevant, what matters is using the same criteria to reshape the results back.

Usually, a physical system could depend on many parameters (*e.g.* inlet velocity $\mathbf{u}$, pressure p, ...), but unless an extensive experimental/numerical campaign is conducted, it is not possible to explore the entire parameter space to fully understand the system's response. Moreover, could be computationally demanding to simulate the physical system for a long period of time. As a consequences, several techniques based on linear algebra methods (and later inherited by machine learning approaches) have been developed to reduce the complexity of physical systems and extrapolating the minimum number of important informations to recreate (and predict) the time/space evolution and/or the system behaviour

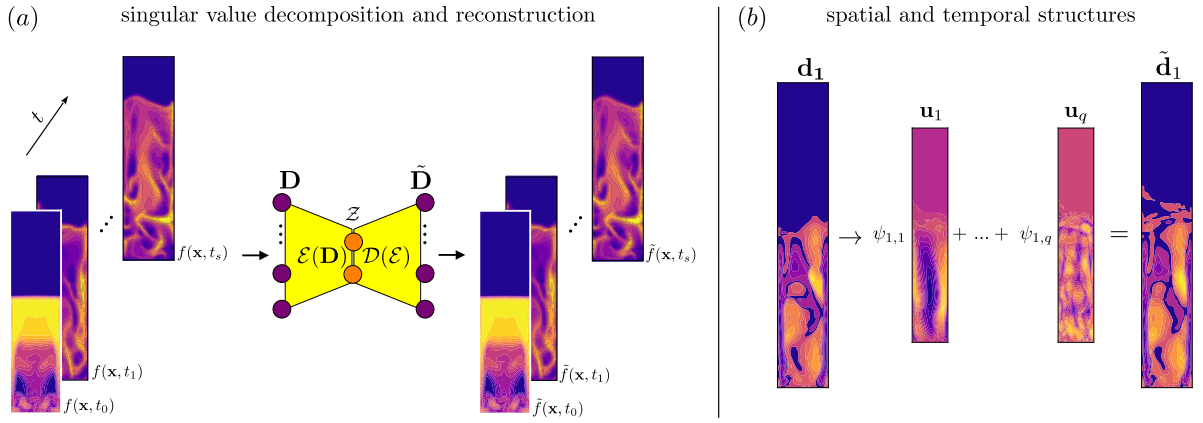

Figure 2: **Data matrix decomposition. (a)** Simple representation of the SVD method. The generic field $f(\mathbf{x}, t)$ (solid phase volume fraction in the figure) is measured at each mesh points for all the s simulated timestep (left part). Each measurements (snapshot on the left part) is used to populate a column of the data matrix $\mathbf{D}$. Subsequently, the data are projected ($\mathcal{E}(\mathbf{D})$) into a reduced-dimensional space $\mathcal{Z}$ and reconstructed through a decoding process $\mathcal{D}(\mathcal{E})$ that achieves an approximation $\overline{\mathbf{D}}$ of the original data (snapshot in the right part of the panel). Note that, in panel (a), reactor height is truncated for figure clarity and compactness. **(b)** Decomposition of a solid fraction snapshot written as a linear combination of the truncated spatial $\mathbf{u}_k$ and temporal ψ_k basis.

for different input parameters. One of the most used techniques, is the dimensionality reduction based on SVD method, here briefly summarized and simplified for the reader convenience.
The central point (among others) of SVD is to reduce and compress a large number of interdependent variables into a smaller of independent variables while preserving as much information and variance from the original variables as possible. Consider the data-set $\mathbf{D}$ described above, populated with s observation of m variables, the SVD provides an approximation of the original data-set using $q \leq \min(s, m)$ variables. This means that, the measurements $\mathbf{d}_k \in \mathbb{R}^m$ is projected (or encoded) into a lower dimensional vector $\psi \in \mathbb{R}^q$.
With the SVD method, we decompose a matrix $\mathbf{D}$ in a product of three matrices

$$\mathbf{D} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \tag{1}$$

where $\mathbf{U} \in \mathbb{R}^{m \times m}$ and $\mathbf{V} \in \mathbb{R}^{s \times s}$ are unitary matrices ($\mathbf{U^T U} = \mathbf{U U^T} = \mathbb{I} = \mathbf{V V^T} = \mathbf{V^T V}$, where $\mathbb{I}$ is the identity matrix) and $\mathbf{\Sigma} = \left[\hat{\mathbf{\Sigma}}, \mathbf{0}\right]^T \in \mathbb{R}^{m \times s}$ with non-negative elements on the main diagonal $\hat{\mathbf{\Sigma}}$ and zeros elsewhere. The $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{\Sigma}$ are called left singular matrix, right singular matrix and singular value matrix, respectively. The decomposition matrices are computed solving the following eigenvalue problems:

$$\mathbf{D}\mathbf{D}^T = \mathbf{U} \begin{bmatrix} \hat{\mathbf{\Sigma}}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{U}^T, \tag{2}$$

$$\mathbf{D}^T\mathbf{D} = \mathbf{V}\hat{\mathbf{\Sigma}}^{\mathbf{2}}\mathbf{V}^T, \tag{3}$$

where $\mathbf{D}\mathbf{D}^T$ and $\mathbf{D}^T\mathbf{D}$ are called spatial and temporal correlation matrices, respectively. Eqs.[2-3] provide a practical and very important interpretation of those matrices: if elements in the $\mathbf{\Sigma}$ diagonal are arranged in descendent order and, at the same time, the columns of $\mathbf{U}$ and rows of $\mathbf{V}$ are moved coherently, the latter describe how much correlation in columns and rows of $\mathbf{D}$ is captured, respectively. Since in our settings the column elements of $\mathbf{D}$ are the spatial measurements of the generic field $f(\mathbf{x}, t)$, the column of $\mathbf{U}$ represent the spatial structure and, similarly, the rows of $\mathbf{V}^T$ are the temporal structure basis. In order to have a reduced representation (or approximation) $\tilde{\mathbf{D}}$ of the data matrix, we retained only q column of $\mathbf{U}$ and rows of $\mathbf{V}^T$, obtaining the truncated representation of $\mathbf{D}$

$$\tilde{\mathbf{D}} \approx \mathbf{U}_q\mathbf{\Sigma}_q\mathbf{V}_q^T, \tag{4}$$

where the subscript q indicates the number of retained vectors. In our numerical set-up, the number m of mesh grid points is larger than the simulated timestep s, hence $q \leq s$.
Expanding the Eq. (4) for the ith column $\mathbf{d}_k$, we obtain that

$$\mathbf{d}_k \approx \sum_{k=1}^{q} \mathbf{u}_j \sigma_j v_{j,k} = \sum_{j=1}^{q} \mathbf{u}_j \psi_{j,k}, \tag{5}$$

where $\mathbf{u}_k$ is the kth column of $\mathbf{U}_q$, σ_k is the kth singular value of $\mathbf{\Sigma}_q$, $v_{j,k}$ is the kth elements of the jth right singular vector and $\psi_{j,k} = \sigma_j v_{j,k}$ that are collected in a matrix $\mathbf{\Psi} \in \mathbb{R}^{q \times s}$. Eq.(5), provide a very simple way to decompose the data matrix (and thus the generic field $f(\mathbf{x}, t)$) as a linear combination of the spatial structures in which the coefficient contain temporal information, see Fig.2(b).
In the following, and in particular in Results and Discussions section, the data matrix is centered before applying the SVD, this means that the mean of all the columns is subtracted from each column:

$$\mathbf{D}_o = \mathbf{D} - \overline{\mathbf{D}}, \tag{6}$$

$$\overline{d}_{k,j} = \frac{1}{s} \sum_{i=0}^{s} d_{k,i}, \tag{7}$$

where $\mathbf{D}_o$ is the centered data matrix, $\overline{\mathbf{D}} \in \mathbb{R}^{m \times s}$ is matrix of centers and $\overline{d}_{k,j}$ is the generic entries of the center matrix. With this transformation, the Eq.(5) becomes

$$\mathbf{d}_{o,k} \approx \overline{\mathbf{d}}_k + \sum_{j=1}^{q} \mathbf{u}_j \psi_{j,k}, \tag{8}$$

where $\overline{\mathbf{d}}_k$ is the kth column vector of $\overline{\mathbf{D}}$ (note that all the $\overline{\mathbf{d}}_k$ are equals).
From here on, the subscript o will be omitted for brevity, and the matrix $\mathbf{D}$ will always refer to its centered counterpart unless otherwise indicated.
As a final comment, the SVD method can be seen as a linear autoencoder, see Fig.2(a). The linear encoder function $\mathcal{E}(\mathbf{D}) : \mathbb{R}^m \to \mathbb{R}^q$ is a projection of the data matrix in a lower dimensional space described by q vectors such that $\mathcal{Z} = \mathbf{U}_q^T \mathbf{D} \in \mathbb{R}^{q \times q}$, where the column of $\mathbf{U}_q$ are the orthogonal basis vectors of that space and, since they contain the most important information in the data matrix, they are usually called principal components (PC) in the machine learning fields. $\mathcal{Z}$ is the compressed form of the original data matrix. The decoder function, is also a linear combination of these basis such that $\mathbf{D} \approx \mathcal{D}(\mathcal{E}) = \tilde{\mathbf{D}} = \mathbf{U}_q \mathcal{Z} = \mathbf{U}_q \mathbf{U}_q^T \mathbf{D}$ and replacing Eq.(1) in the last equation we obtain that $\mathcal{Z} = \mathbf{\Sigma}_q \mathbf{V}_q^T$, and the rows of $\mathcal{Z}$ (*i.e.* ψ_k) are called PC scores. It is worth noting that while $\mathbf{U}_q^T \mathbf{U}_q = \mathbb{I}$, it is no longer true that $\mathbf{U}_q \mathbf{U}_q^T = \mathbb{I}$, this means that, to achieve a good approximation of the data matrix, the latter relation must be as close as possible to the identity. For further details regarding the SVD and principal component analysis, the interested reader is referred to [24, 25, 27, 36].
In principle, in the data matrix different fields can be stacked to compose a column $\mathbf{d}_k$. In our work, we used the SVD for each fields separately to obtain a better representation for each physical quantities namely, velocity of the gas carrier, pressure,and solid volume fraction of the bed. Moreover, in order to obtain a good approximation of the data-set, the number of the retained principal directions (column of $\mathbf{U}_q$), was chosen looking at the singular values. In particular, we retained those principal directions whose the corresponding singular values capture at least the 95% of the variance. Therefore, the number of PCs differ for each quantity of interest.

2.3 Gaussian Processes

The accurate prediction of the quantity of interest for unexplored points of the parameter space (*i.e* nitrogen velocity $\mathbf{u}_{in,N_2}$), requires an accurate estimation of the PC and PC score for those points. Gaussian Processes are a powerful tools for regression problems, based on a probabilistic description. GP regression models assume that the function (distribution) relating the inputs to the outputs is drawn from a Gaussian distribution with specified mean and covariance functions. Usually, the prior mean of the target distribution is supposed to be zero and the covariance function are also called kernel function. The kernel functions are typically chosen in order to increase the correlation for nearby points compared to dissimilar points. In this work the squared exponential kernel function is used [37].
Finally, the matrices are interpolated using the nitrogen velocity as parameter and interpolation weight. A detailed discussion of GP and in general for kernel methods for regression, can be found in [24, 38].

3 Results and Discussions

In this section the results of this study are presented. First, a discussion about the SVD application and results is presented, followed by the results on the regression processes on the compressed representation of the data matrix.

3.1 SVD e dimensionality reduction

The SVD method and the dimensionality reduction was conducted on a data matrix with $s = 270$ observation (columns), that correspond to an observation every 0.1 s for 27 s, for each quantity of interest, namely $\mathbf{u}_{N_2}, p$ and v. To find the number of modes to retain, the so called explained variance of the singular values is used. Given the ordered singular values, elements on the $\mathbf{\Sigma}$ diagonal, from the data matrix $\mathbf{D}$ decomposition, they are normalized with respect to the sum over the singular values

$$\sigma_j = \frac{\hat{\sigma}_j}{\sum_{k=1}^{s} \sigma_k}, \tag{9}$$

where $\hat{\sigma}_k$ is a singular value and σ_k is the normalized one. In the Methods section, we have defined the retained number of modes as the minimum number of modes such that the captured variance is $=> 95\%$, *i.e.* the cumulative variance, also called explained variance. In Fig. 3(a-c), three examples of the number of the retained singular values and the explained variance are reported. In particular, in the Fig. 3 some illustrative results of the fluidized-bed dimensionality reduction for $\mathbf{u}_{in,N_2} = 0.20\,\mathrm{m/s}$ and $t = 1.5\,\mathrm{s}$ are reported. As can be observed, for the pressure field five modes are enough to reconstruct the field for all the time considered while maintaining the error low, see bottom part of Fig. 3(a). For more complex field, such as velocity and volume solid fraction, more modes are required to recover the fields from the reduced representation of the data matrix. In fact, for the nitrogen velocity and solid volume fraction 20 and 33 modes are needed, respectively, but despite the increased complexity, the recovery error remains low. Note that the number of preserved modes q is the dimension of the compressed space where $\mathbf{D}$ is projected. In our settings, we started with a data matrix $(m, s) = (60 \times 10^3, 280)$, subsequently transformed into a matrix of shape (q, q). For completeness in Fig. 3(d-f) a comparison between the observed fields and the recovered ones is reported. Finally, in Fig. 3(g-i), the R^2 values of the reconstruction for different values of the nitrogen inlet velocity evaluated for each time step are shown. The R^2 value, is grater than 0.94 for every predicted field, confirming the good recovery from the chosen modes. Moreover, as a general trend, can be observed that the error increases as the $\mathbf{u}_{in,N_2}$ increases.

3.2 Field prediction and regression

In fluidized-bed reactors, the inlet velocity of the carrier gas could be divided in two categories: when it is greater than or when it is less than the minimum fluidization velocity u_{mf}. At low gas velocities the bed particles remain packed, while as the gas velocity increases, the particles start to move, the bed expands and the reactor enters in the fluidized region. To ensure good performance of the reactors, the inlet gas velocity must be greater than the $u_{m,f}$. Based on this working principle of fluidized-beds, in this study we focused the effort to predict the quantities of interest (pressure p, nitrogen velocity $\mathbf{u}_{N_2}$ and solid volume fraction v) for inlet nitrogen velocity $\mathbf{u}_{in,N_2} > u_{mf} = 0.06$ m/s for the considered reactor set-up. In particular, a set of simulations are performed to train the ROM considering eight values of the N_2 inlet velocity (observations) $u_{in,N_2} = 0.06$, 0.08, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35 m/s. In order to test the model for different $\mathbf{u}_{in,N_2}$, the training set was populated removing alternatively one of the observation, and testing the ROM with the removed velocity. Therefore, the training phase of the GP is repeated each time an observation is removed, for a total of 8 training-testing phases.

In this section, hereinafter, the term *reconstruction* is used to indicate the data matrix reconstruction through its compressed form achieved by SVD, while the term *prediction* refers to the field reconstruction for $\mathbf{u}_{in,N_2}$ not used during the training phase.

As mentioned above in the Methods section, the regression process is applied, for each observation, on the PC matrix $\mathbf{U}_q$ and on PC scores matrix $\mathcal{Z}$. The predicted field is then obtained reconstructing it from its decomposition matrices approximation. In Fig. 4, the results obtained from the prediction of the reactor behaviour at $t = 27$ s for $u_{in,N_2} = 0.2$ m/s and therefore not using this velocity during the training, for the three fields of interest are reported. In each panels the observation against the prediction results (colored dots) are presented together with three red solid-lines: the external lines represent a prediction error of 10%, while the middle one corresponds to a perfect prediction. For the pressure field, Fig. 4(a), can be observed how the $\mathbf{U}_q$ and $\mathcal{Z}$ are well estimated generally maintaining the error below the 10% limit. This low error propagates during the reconstruction leading to a greater

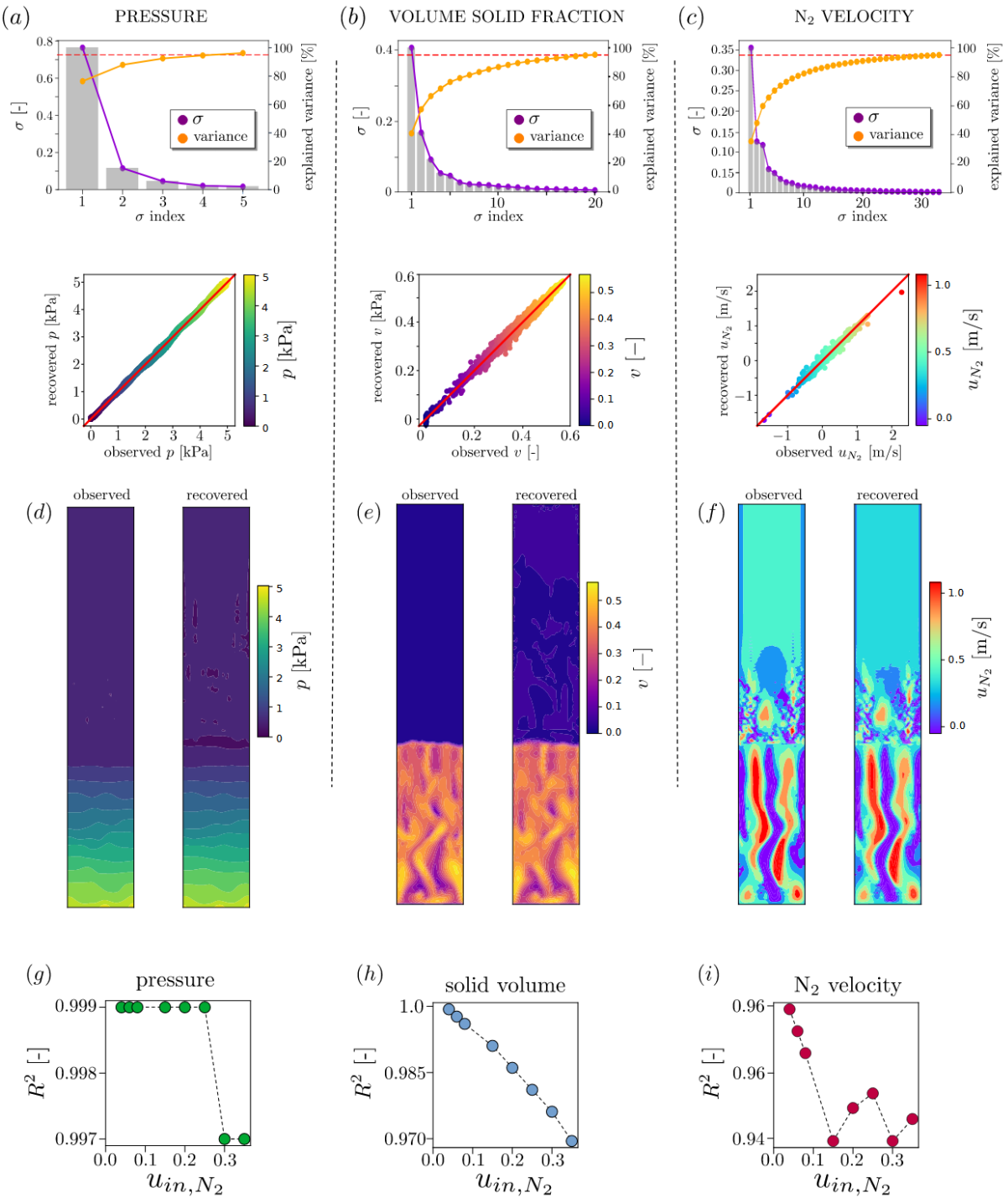

Figure 3: **SVD and reconstruction results.** Results on SVD decomposition and reconstruction for a fluidized-bed reactors for $\mathbf{u}_{in,N_2} = 0.2$ m/s and $t = 1.5s$ are reported. **(a-c)** In the top part the explained variance and the retained singular values σ for the pressure, volume solid fraction and N_2 velocity fields are reported. The violet solid line and grey bars represent the normalized singular values according Eq. (9). The orange line described the evolution of the explained (cumulative) variance. While in bottom part, the quality of the recovered field are presented using parity plots. **(d-f)** Graphical comparison between the recovered and observed fields. **(g-i)** Error of the reconstruction for all the inlet velocity and time-step.

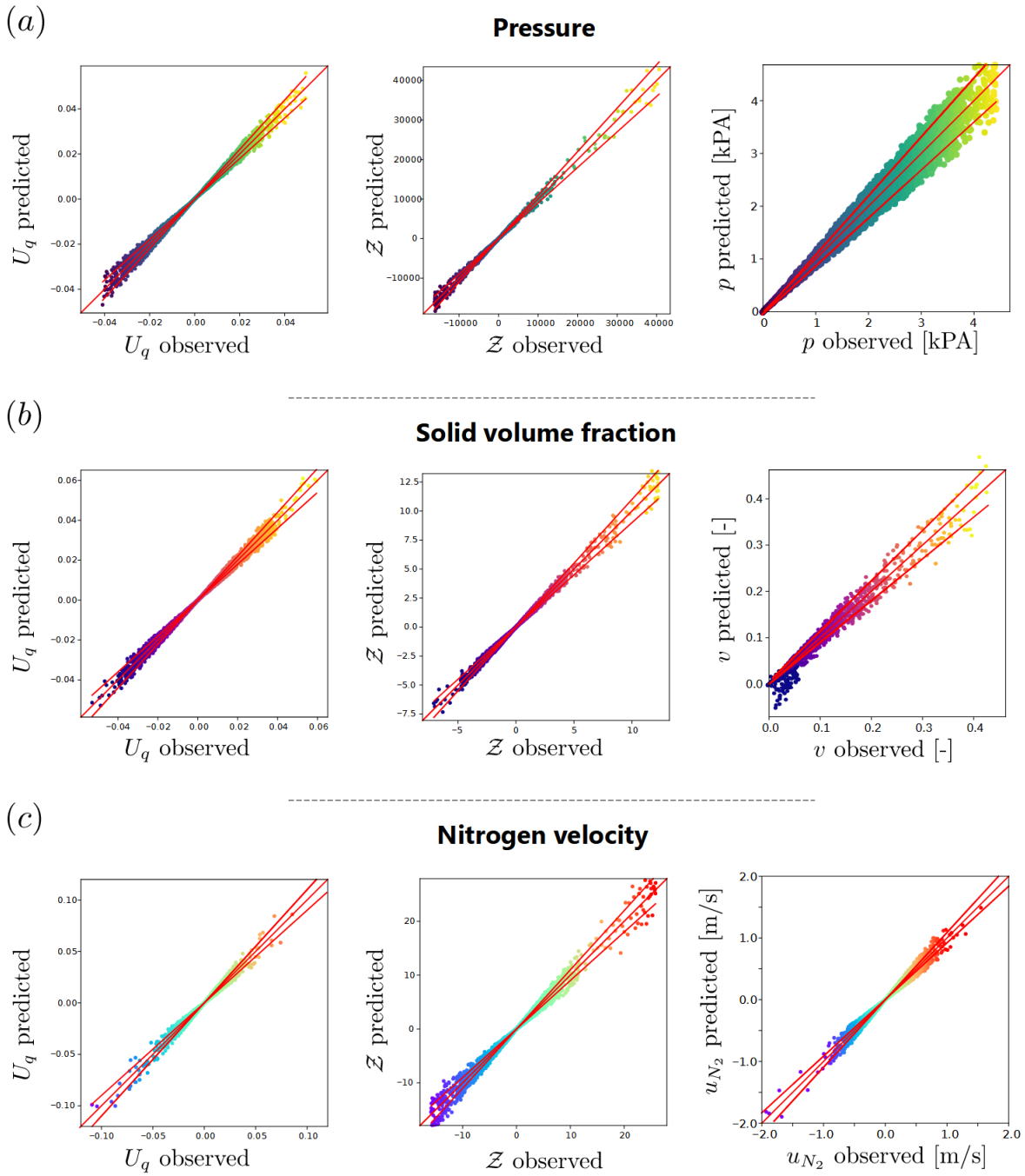

Figure 4: **Comparison between prediction and observation.** Parity plots for quantification of the quality of the prediction of the regression. In particular, results for $\mathbf{u}_{in,N_2}$ and $t = 27\,\mathrm{s}$ are reported. The colored dots describe the comparison between prediction and observation. The red solid lines represent the perfect prediction (central line), and $\pm 10\%$ error bounds of the prediction (external lines).

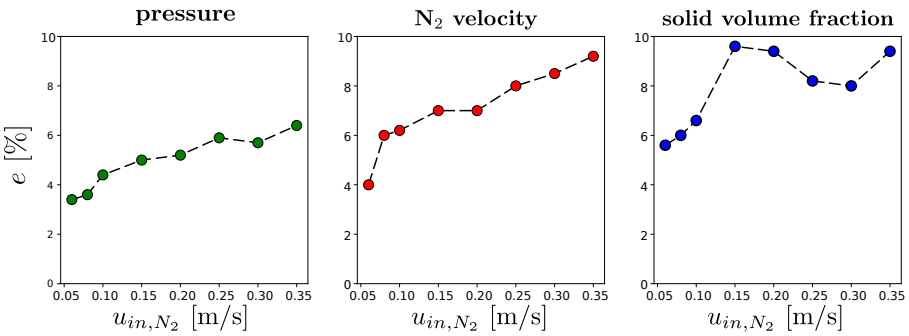

Figure 5: **Overall prediction error**. Total prediction error evaluated according the Eq. (10), for the pressure p (green circles), N_2 velocity u_{N_2} (red circles) and solid volume fraction v (blue circles).

error during the field reconstruction. Even with the uncertainty propagation, the prediction error remains low, although it does increase slightly as pressure increases. Another important results is related to the pressure values. In fact, as highlighted in the Methods section, the pressure here reported is not an absolute pressure, but is a relative pressure with respect to the outlet one. Therefore, since the gas carrier must flow from the bottom to the top of the reactor, this pressure must be positive. During the SVD (or PCA/POD) decomposition and reconstruction, the positivity of the pressure is not guaranteed. In fact, usually, in steady state problems to ensure that some physics properties are respected, some constraints can be added during the SVD/PCA approach [13]. In fact, as shown in the right side of Fig. 4(b), the problem of unfeasible values of v are predicted. For low values of the observed v, the predicted counterpart assumes negative value and this is obviously not possible in the real world, since a volume fraction can not be negative by definition. Moreover, in this work, this issue arises from the fact that the prediction (of an highly unsteady flow) is obtained from the regression over the decomposition matrices and not directly over the volume fraction field [39, 40]. In the light of this considerations, it is less straightforward to define constraints on the elements of the matrices $\mathbf{U}_q$ and $\mathcal{Z}$. While the mathematical interpretation of the matrices can be established, it is less clear how such constraints can be reflected in the actual quantitative values of these elements. Despite the negative values predicted for small values of v, the global error is generally contained within the limit of 10%. For $\mathbf{u}_{N_2}$, similar trends can be observed, where a good satisfactory level of accuracy is also achieved as depicted by the parity plots in Fig. 4(c).
In Fig. 5, the time averaged prediction error for each predicted field and for all the velocity is reported. The error e is evaluated bthrough the following equation

$$e = \frac{1}{(\max_{i,j}\{\mathbf{D}\} - \min_{i,j}\{\mathbf{D}\})t} \sum_{j=1}^{t}\sum_{i=1}^{m} |d_{i,j} - \hat{d}_{i,j}|, \tag{10}$$

where $\mathbf{D}$ is the data matrix, t is the number of snapshots in the data matrix, m is the number of grid points, $\mathbf{d}$ and $\hat{\mathbf{d}}$ are the column data matrix $\mathbf{D}$ and its prediction $\hat{\mathbf{D}}$, respectively. As shown in Fig. 5, the error increases as the velocity increases, a trend observed for all fields with some exceptions regarding the solid fraction v. This behaviour can be attributed to the chaotic motion of the bed material inside the reactor. In fact, as the inlet velocity increases, the bed becomes more fluidized leading to less accurate prediction, although the averaged error remain smaller than 10%.
Predicting the time evolution of solid volume fraction and pressure drop within the reactor, allows for controlling mass transport, mass transfer, and heat transfer between the involved species. Such control, acting on the nitrogen inlet velocity, can alter the behaviour of the fluidized bed, steering the process towards the production of specific outputs (*e.g* gas or oil) [41]. Moreover, knowing the solid fraction distribution helps optimizing the mixing of particles and gasses, which is essential for achieving uniform temperature. In perspective, the knowledge of physical behaviours allows also the evaluation of the fluidization quality, to quantify the efficiency of the heat transfer and, consequently, of chemical reaction control.

4 Conclusions

Pyrolysis is a reliable thermochemical process for energy and chemicals recovery. Pyrolytic reactor are usually designed to process limited classes of feedstock or to obtain a certain output. Development of advanced control systems could help to improve these limitation, but a comprehensive knowledge of the physical phenomena that occurs within the reactors varying the operating conditions, is needed. The present work, is intended to be a preliminary study aiming to demonstrate the capability of data-driven ROM to predict the responses of very complex systems (like fluidized-bed reactor), even for operating condition not explored during the training. Having an accurate ROM that can reliably and instantaneously predict the behaviour of these systems, enables the possibility to develop real-time controller to increase the pyrolysis flexibility.
Here, we proposed a data-driven approach based on dimensionality reduction (SVD) and data regressions through GP to predict the behavior of a fluidized-bed reactor for pyrolysis. In particular, pressure, nitrogen velocity (fluidising gas) and solid volume fraction fields are predicted for different nitrogen inlet velocity. The prediction and control of these quantities through the inlet velocity of the fluidising gas, could help to improve the mixing and therefore maximizing the mass and heat transfer efficiency. The developed ROM is able to predict these fields with an error $<10\%$ for all the simulated time, demonstrating the applicability of this approach in real-time control applications. In future works, we will extend and improve the methodology increasing the number of input parameters and introducing chemical reactions for prediction of products for different operating conditions.

Acknowledgements
The authors acknowledge support from Project ECS 0000024 Rome Technopole - CUP: E83C22003240001, NRP Mission 4 Component 2 Investment 1.5, Funded by the European Union – NextGenerationEU.

References

[1] Sonil Nanda and Franco Berruti. A technical review of bioenergy and resource recovery from municipal solid waste. Journal of hazardous materials, 403:123970, 2021.

[2] R Santagata, M Ripa, A Genovese, and S Ulgiati. Food waste recovery pathways: Challenges and opportunities for an emerging bio-based circular economy. a systematic review and an assessment. Journal of Cleaner Production, 286:125490, 2021.

[3] Dezhen Chen, Lijie Yin, Huan Wang, and Pinjing He. Pyrolysis technologies for municipal solid waste: a review. Waste management, 34(12):2466–2486, 2014.

[4] Matteo Baldelli, Lorenzo Bartolucci, Stefano Cordiner, Giorgio D'Andrea, Emanuele De Maina, and Vincenzo Mulone. Biomass to h2: Evaluation of the impact of pv and tes power supply on the performance of an integrated bio-thermo-chemical upgrading process for wet residual biomass. Energies, 16(7):2966, 2023.

[5] Sascha RA Kersten, Xiaoquan Wang, Wolter Prins, and Wim PM van Swaaij. Biomass pyrolysis in a fluidized bed reactor. part 1: Literature review and model simulations. Industrial & engineering chemistry research, 44(23):8773–8785, 2005.

[6] KN Yogalakshmi, P Sivashanmugam, S Kavitha, Yukesh Kannah, Sunita Varjani, S AdishKumar, Gopalakrishnan Kumar, et al. Lignocellulosic biomass-based pyrolysis: A comprehensive review. Chemosphere, 286:131824, 2022.

[7] Tamer YA Fahmy, Yehia Fahmy, Fardous Mobarak, Mohamed El-Sakhawy, and Ragab E Abou-Zeid. Biomass pyrolysis: past, present, and future. Environment, Development and Sustainability, 22:17–32, 2020.

[8] Muhammad Saad Qureshi, Anja Oasmaa, Hanna Pihkola, Ivan Deviatkin, Anna Tenhunen, Juha Mannila, Hannu Minkkinen, Maija Pohjakallio, and Jutta Laine-Ylijoki. Pyrolysis of plastic waste: Opportunities and challenges. Journal of Analytical and Applied Pyrolysis, 152:104804, 2020.

[9] Steven L Brunton, Bernd R Noack, and Petros Koumoutsakos. Machine learning for fluid mechanics. Annual review of fluid mechanics, 52(1):477–508, 2020.

[10] Benjamin Schleich, Nabil Anwer, Luc Mathieu, and Sandro Wartzack. Shaping the digital twin for design and production engineering. CIRP annals, 66(1):141–144, 2017.

[11] Mengnan Liu, Shuiliang Fang, Huiyue Dong, and Cunzhi Xu. Review of digital twin about concepts, technologies, and industrial applications. Journal of manufacturing systems, 58:346–361, 2021.

[12] Thomas H-J Uhlemann, Christoph Schock, Christian Lehmann, Stefan Freiberger, and Rolf Steinhilper. The digital twin: demonstrating the potential of real time data acquisition in production systems. Procedia Manufacturing, 9:113–120, 2017.

[13] Gianmarco Aversano, Aurélie Bellemans, Zhiyi Li, Axel Coussement, Olivier Gicquel, and Alessandro Parente. Application of reduced-order models based on pca & kriging for the development of digital twins of reacting flow applications. Computers & chemical engineering, 121:422–441, 2019.

[14] Nedunchezhian Swaminathan and Alessandro Parente. Machine Learning and Its Application to Reacting Flows: ML and Combustion. Springer Nature, 2023.

[15] Li-Tao Zhu, Xi-Zhong Chen, Bo Ouyang, Wei-Cheng Yan, He Lei, Zhe Chen, and Zheng-Hong Luo. Review of machine learning for hydrodynamics, transport, and reactions in multiphase flows and reactors. Industrial & Engineering Chemistry Research, 61(28):9901–9949, 2022.

[16] Hao Luo, Xiaobao Wang, Xinyan Liu, Lan Yi, Xiaoqin Wu, Xi Yu, Yi Ouyang, Weifeng Liu, and Qingang Xiong. Machine learning based prediction of biomass pyrolysis with detailed reaction kinetics for thermally-thick particles: from 1d to 0d. Chemical Engineering Science, 280:119060, 2023.

[17] Maria Puig-Arnavat, J Alfredo Hernández, Joan Carles Bruno, and Alberto Coronas. Artificial neural network models for biomass gasification in fluidized bed gasifiers. Biomass and bioenergy, 49:279–289, 2013.

[18] Ankun Xu, Huimin Chang, Yingjie Xu, Rong Li, Xiang Li, and Yan Zhao. Applying artificial neural networks (anns) to solve solid waste-related issues: A critical review. Waste Management, 124:385–402, 2021.

[19] Daniel Serrano, Iman Golpour, and Sergio Sánchez-Delgado. Predicting the effect of bed materials in bubbling fluidized bed gasification using artificial neural networks (anns) modeling approach. Fuel, 266:117021, 2020.

[20] Blake R Hough, David AC Beck, Daniel T Schwartz, and Jim Pfaendtner. Application of machine learning to pyrolysis reaction networks: Reducing model solution time to enable process optimization. Computers & Chemical Engineering, 104:56–63, 2017.

[21] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. Nature Reviews Physics, 3(6):422–440, 2021.

[22] Aref Ghaderi, Ramin Akbari, Yang Chen, and Roozbeh Dargazany. A knowledge-driven physics-informed neural network model; pyrolysis and ablation of polymers. arXiv preprint arXiv:2209.11749, 2022.

[23] Weiqi Ji, Franz Richter, Michael J Gollner, and Sili Deng. Autonomous kinetic modeling of biomass pyrolysis using chemical reaction neural networks. Combustion and Flame, 240:111992, 2022.

[24] Christopher M Bishop and Nasser M Nasrabadi. Pattern recognition and machine learning, volume 4. Springer, 2006.

[25] Ian T Jolliffe. Principal component analysis for special types of data. Springer, 2002.

[26] Shai Shalev-Shwartz and Shai Ben-David. Understanding machine learning: From theory to algorithms. Cambridge university press, 2014.

[27] SL Brunton and JN Kutz. Data driven science & engineering-machine learning, dynamical systems, and control (p. 572). databook. uw. edu. 2017.

[28] S Walton, O Hassan, and K Morgan. Reduced order modelling for unsteady fluid flow using proper orthogonal decomposition and radial basis functions. Applied Mathematical Modelling, 37(20-21):8930–8945, 2013.

[29] Panneerselvam Ranganathan and Sai Gu. Computational fluid dynamics modelling of biomass fast pyrolysis in fluidised bed reactors, focusing different kinetic schemes. Bioresource technology, 213:333–341, 2016.

[30] Madhava Syamlal, Thomas J O'Brien, et al. Computer simulation of bubbles in a fluidized bed. In AIChE Symp. Ser, volume 85, pages 22–31. Publ by AIChE, 1989.

[31] DJ Gunn. Transfer of heat or mass to particles in fixed and fluidised beds. International Journal of Heat and Mass Transfer, 21(4):467–476, 1978.

[32] Cli KK Lun, Stuart B Savage, DJ Jeffrey, and Nicholas Chepurniy. Kinetic theories for granular flow: inelastic particles in couette flow and slightly inelastic particles in a general flowfield. Journal of fluid mechanics, 140:223–256, 1984.

[33] Mo Syamlal. The particle-particle drag term in a multiparticle model of fluidization. Technical report, EG and G Washington Analytical Services Center, Inc., Morgantown, WV (USA), 1987.

[34] Ansys Inc. ANSYS® Academic Research CFD, Release 2024 R1, Fluent User's Guide. 2024.

[35] Ansys Inc. ANSYS® Academic Research CFD, Release 2024 R1, Fluent Theory Guide. 2024.

[36] MA Mendez, M Balabane, and J-M Buchlin. Multi-scale proper orthogonal decomposition of complex fluid flows. Journal of Fluid Mechanics, 870:988–1036, 2019.

[37] Matthias Seeger. Gaussian processes for machine learning. International journal of neural systems, 14(02):69–106, 2004.

[38] Christopher Williams and Carl Rasmussen. Gaussian processes for regression. Advances in neural information processing systems, 8, 1995.

[39] James C Sutherland and Alessandro Parente. Combustion modeling using principal component analysis. Proceedings of the Combustion Institute, 32(1):1563–1570, 2009.

[40] Chinchun Ooi, Quang Tuyen Le, My Ha Dao, Van Bo Nguyen, Hoang Huy Nguyen, and Te Ba. Modeling transient fluid simulations with proper orthogonal decomposition and machine learning. International Journal for Numerical Methods in Fluids, 93(2):396–410, 2021.

[41] Daizo Kunii and Octave Levenspiel. Fluidization engineering. Butterworth-Heinemann, 1991.