HUNGARIAN JOURNAL OF
INDUSTRY AND CHEMISTRY
HJIC

DEVELOPMENT OF COMPARTMENT MODELS FOR DIAGNOSTIC PURPOSES

BÁLINT LEVENTE TARCSAY[*1], ÁGNES BÁRKÁNYI[1], TIBOR CHOVÁN[1], AND SÁNDOR NÉMETH[1]

[1]Research Centre for Biochemical, Environmental and Chemical Engineering, University of Pannonia, Egyetem u. 10, Veszprém, 8200, HUNGARY

The importance of recognizing the presence of process faults and resolving these faults is continuously increasing parallel to the development of industrial processes. Fault detection methods which are both robust and sensitive help to recognize the presence of faults in time to avoid malfunctions, financial loss, environmental damage or loss of human life. In the literature, the use of various model-based fault detection methods has gained a considerable degree of popularity. Methods usually based on black-box models, data-based techniques or models using symbolic logic, e.g. expert systems, have become widespread. White-box models, on the other hand, have been applied less despite their considerable robustness because of multiple reasons. Firstly, their complexity and the relatively vast amount of technological and modelling knowledge needed to construct them for industrial systems. Secondly, their large computational demand which makes them less suitable for online fault detection. In this study, the aim was to resolve these problems by developing a method to simplify the complex Computational Fluid Dynamics models employed to describe the equipment used in the chemical industry into less complex model structures. These simpler structures are Compartment Models, a type of white-box model which breaks down a complex system into smaller units with idealized behaviour. In the case of a small number of compartments, the computational load of such models is not significant, therefore, they can be employed for the purposes of online fault detection while providing an accurate representation of the system. For the purpose of identifying the compartmental structure, fuzzy logic was employed to create a model which approximates the real behaviour of the system as accurately as possible. Our future objective is to explore the possibility of combining this model with various diagnostic methods (expert systems, Bayesian networks, parity relations, etc.) and derive robust tools for the purpose of fault detection.

Keywords: Compartment Model, Computational Fluid Dynamics, expert system, fuzzy logic

1. Introduction

Fault detection and isolation has become a cardinal problem in industrial practices. As the intricacy of technological processes increases, the probability of faults arising also grows. However, due to the stricter demands for environmentally friendly and profitable technologies, the presence of faults which could potentially cause financial loss, environmental hazards or loss of life cannot be tolerated. Faults such as those defined by Venkatasubramanian et al. [1] describe the consequences when an observed process variable deviates from its expected value. The source of such process faults is known as a root cause. Naturally, if left unmanaged, faults can propagate and lead to serious disasters, e.g., the Bhopal disaster or the Texas City Refinery explosion [2, 3]. To circumvent such occurrences, modern knowledge-based fault detection techniques have become commonplace.

These methods use the same common principle for fault detection. Firstly, some knowledge is used to create a model of the system under normal operating conditions. Then the actual operation of the system is observed and, if the behaviour of the system does not match the assumptions drawn from the model, then an abnormality is present [4]. After identifying the presence of an abnormality, various approaches using different techniques trace the fault signature back to the root cause depending on the type of knowledge used to create the system model. Therefore, it is possible to differentiate between knowledge that originates from quantitative as well as qualitative models and that from relations based on process history [1, 5, 6].

The aforementioned categorization of knowledge used for fault detection methods stems from the early works of Venkatasubramanian [1] and is arbitrary. The expectations of fault detection and isolation methods are that they should be robust, sensitive and accurate. Robust, in this sense, means that the method should work reliably, even in the presence of noise, disturbances and changes in operating conditions. Sensitive and accurate means that the method should even be able to recognize small pro-

*Correspondence: tarcsayb@fmt.uni-pannon.hu

cess faults and reliably identify their root causes. Usually, the bigger issue with most fault detection methods is their robustness. Commonly used models for the purposes of fault detection, e.g., a posteriori models, neural networks and statistical methods, all employ arbitrary relationships between system variables which only hold true in a certain operating regime [4].

If the system itself or the operating point changes, these models often become unreliable. Therefore, they lack both robustness and flexibility. Among quantitative models used for fault detection, another type of model, the a priori or white-box model, is listed. This modelling technique uses the laws of physics, chemistry or any applicable scientific discipline to describe the behaviour of various systems. Since the laws of nature cannot be violated, these models are incredibly robust and hold true for a system regardless of its operating point.

The problem is that these models are complex and difficult to create for industrial systems since they require a vast amount of technological and modelling knowledge. Additionally, they often require a significant computational load to be solved that renders them impractical for the purposes of online fault detection [1]. This is why they have not garnered widespread use in the field of diagnostic procedures.

Over the course of this work, our goal was to propose a method for simplifying the complex Computational Fluid Dynamics (CFD) of white-box modelling techniques to describe the behaviour of units used in the chemical industry. CFD methods have become widespread over the last few decades for the purpose of solving the material and energy balance equations of complex systems to describe transport phenomena as well as evaluate flow patterns, velocity fields and mixing in fluid systems. These numerical methods involve the discretization of the differential equations that describe the behaviour of the unit as well as solving them through direct or iterative computation methods [7].

An alternative to these methods is the use of Compartment Models (CM). CMs emulate the behaviour of complex systems by identifying individual elements of the system that can be approximated using idealized models [8]. In the case of chemical technologies, the flow and transport phenomena in a system can be approximated by breaking down the system into compartments that exhibit idealized flow behaviour. CMs have become widespread in pharmacokinetic experiments to simulate drug transport in biological systems as well as during the evaluation and modelling of bioreactors and chemical systems. Savic et al. [9] used a transit CM in pharmacokinetic studies involving the absorption of four different substances. They compared the results to those of a lag time model and concluded that the proposed CM produced a significantly better match to the observed experimental data and can be used as an alternative to the lag time model. Pigou and Morchain [10] employed a CM combined with a population balance model and metabolic model to estimate the behaviour of a microbial population within industrial

bioreactors. Their approach was validated by applying it to a bioreactor containing the *Escherichia coli.* species using experimental data. Spann et al. [11] created a risk assessment methodology for the cultivation of lactic acid bacteria based on a CM. They combined a CM with a biochemical model to monitor the species *Streptococcus thermophilus* online. Using Monte Carlo methods, they evaluated the risk of the biomass size not achieving its desired extent under the operating conditions of the system with regard to uncertainties in the production process. Kaur et al. [12] developed a model for the simulation of a top spray fluidized bed granulator. In their work, they defined two compartments within the system, a "wet" compartment in which aggregation of the particles is dominant and a "dry" compartment in which the breaking up of the particles is mainly observed. By combining the flow model with a population balance model, they managed to observe the effects of operational parameters on the particle size distribution of the granulated product.

The examples show that the CM is a contemporary and popular approach to system modelling. The main question, however, when developing a CM is how to select the appropriate number and characteristics of as well as connections between compartments to obtain a model that accurately represents the flow patterns of the real system. The CM usually employs two types of idealized compartments to approximate a velocity field in a system. These are compartments which are assumed to be perfectly stirred and homogeneous (Continuous Stirred Tank Reactor (CSTR) model) or compartments exhibiting turbulent plug flow (Plug Flow Reactor (PFR) model). These compartments are linked through mixer (M) and distributor (D) models [13]. These latter units are usually only of mathematical significance and their volume is assumed to be zero. The component balance model for the units is showcased in the following equations.

For unit CSTR:

$$\frac{\mathrm{d}c_{\mathrm{CSTR}}(t)}{\mathrm{d}t} = \frac{F_{\mathrm{in}}}{V_{\mathrm{CSTR}}}\left(c_{\mathrm{in}}(t) - c_{\mathrm{CSTR}}(t)\right) \quad (1)$$

$$F_{\mathrm{CSTR}} = F_{\mathrm{in}} \quad (2)$$

For unit PFR:

$$\frac{\partial c_{\mathrm{PFR}}(t,x)}{\partial t} = -v_{\mathrm{PFR}}\frac{\partial c_{\mathrm{PFR}}(t,x)}{\partial x} \quad (3)$$

$$F_{\mathrm{PFR}} = F_{\mathrm{in}} \quad (4)$$

For unit M:

$$c_{\mathrm{M}} = \frac{\sum_{i=1}^{n_{\mathrm{in}}} c_{\mathrm{in},i} F_{\mathrm{in},i}}{\sum_{i=1}^{n_{\mathrm{in}}} F_{\mathrm{in},i}} \quad (5)$$

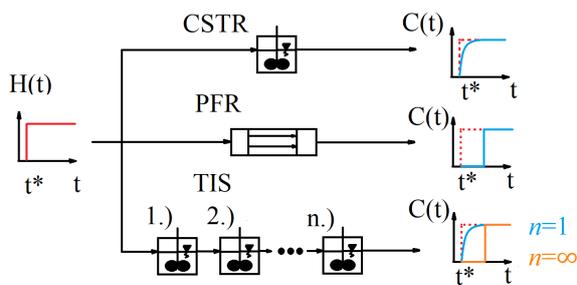$$F_{\mathrm{M}} = \sum_{i=1}^{n_{\mathrm{in}}} F_{\mathrm{in},i} \quad (6)$$

Figure 1: Step responses of ideal flow models [14]

For unit D:

$$\vec{c}_D = \vec{\alpha} c_{in} \tag{7}$$

$$\vec{F}_D = \vec{\alpha} F_{in} \tag{8}$$

The volumetric flow rate (F) entering and exiting the units is also provided by assuming a constant temperature and dilute solutions. The models establish relationships between the volume of the unit (V), the inlet velocity (v) and the inlet as well as actual concentrations in each idealized unit (c). The concentration is introduced as a function of time (t) and/or Cartesian coordinates (x). In the case of the distributor unit, the outlet concentration and flow rate depend on the division rate (α). The idealized CSTR and PFR models exhibit characteristic responses to known input signals such as the Heaviside step function (H) or the Dirac delta function (δ). The step response (C) of the ideal flow models as well as the serial combination of CSTR units in the Tanks in Series (TIS) model are shown in Fig.1.

The response functions can be used to characterize the residence time distribution (RTD) of a chemical species within the system [14]. Therefore, a popular identification method is to obtain the step response of the system and use the RTD function as a basis for identifying the structure of a compartment and the connections between the various compartments. Claudel et al. [15] employed the discipline of possibility theory to provide a method for estimating the structure of a compartment in a unit based on the RTD of the unit and the physical description of the system. They introduced "possibility" and "necessity" rules as well as weighing factors for these rules to assess the RTD function. They summarized the results of the rules to propose the structure of a compartment based on the characteristics of the RTD function.

Egedy et al. [13] used a qualitative approach to identify the structures of compartments in various systems. Their developed algorithm utilized qualitative indicators of the RTD function in an identification algorithm to propose and filter out various CM structures based on their fitness. Approaches which use the CFD model as a basis for the estimation of the CM structure have also become popular in recent studies. Delafosse et al. [16] utilized CFD methods to create a CM structure for approximating the flow characteristics of a bioreactor. Fogarasi et al. [17] developed the CM of a copper leaching reactor based on

experimental and CFD results. Using the CM, they optimized the operational parameters of the leaching reactor. Nauha et al. [18] used a hybrid CFD/CM approach to investigate industrial bioreactors and predict the effects of equipment scale-up on mixing properties within the unit. Weber et al. [19] investigated the use of a hybrid CFD-based CM model for analysing a multiphase loop-reactor. They used the CM model to estimate the mass transfer and drop size distribution within the liquid-liquid extraction part of the loop reactor.

While all of these publications contained significant contributions towards the development of the CFD/CM hybrid model approach, they mostly suffer from the fact that the compartmentalization was conducted manually. Algorithms to automatize the identification of CM structures from CFD have been proposed before. Bezzo et al. [8, 20] used spatial partitioning in their works to develop an algorithm that is capable of aggregating volumes in a unit into compartments and estimating the cross flow between the compartments based on CFD results. Tajsoleiman et al. [21] utilized a similar methodology for the classification of zones within a bioreactor that, based on CFD results, yielded distinctive flow patterns and clustered the results into a CM. Nørregaard et al. [22] developed an algorithm to identify CM structures based on CFD results by using hypothesis-driven logic. They evaluated circumferential, axial and radial bulk flows at different locations within a continuous stirred tank reactor. Using the obtained flow profile, they developed the CM model which resulted in a structure with 56 compartments.

Over the course of this work, the goal was to develop an algorithm capable of identifying the CM structure of a system based on results obtained from CFD methods. The algorithm utilizes two steps to first identify the CM structure and then the flow rates between the individual compartments. Expectations for the algorithm were

- Production of compartments that accurately correspond to the investigated volumes of the physical system

- A small number of compartments to make the CM practical for online computations

- Possibility to integrate empirical knowledge about the system into the compartmentalization process

In future works, the possibility of using the acquired CM as a basis for a fault diagnostic system will be explored.

2. Modelling concept

The proposed algorithm uses CFD results concerning the velocity field within a unit as a basis for identifying the structure of CMs. After partitioning the unit into numerous smaller cells, the algorithm utilizes fuzzy logic to evaluate the local velocity field within the cells and categorize them into one of the idealized models. Adjacent
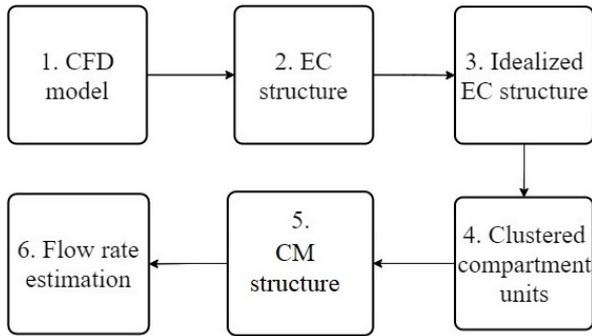
Figure 2: The proposed algorithm

Figure 3: Top and side views of the investigated system

cells that exhibit similar flow behaviour are agglomerated to form compartments showing idealized behaviour. The flow rate between the individual compartments is estimated through optimization. The chart of the proposed algorithm is displayed in Fig. 2.

To summarize the procedure, the compartmentalization process is conducted over six steps.

1. Calculation of a reliable steady state estimate for the velocity field of the system in question using CFD methods.

2. Definition of elementary cells (EC) by dividing the investigated volume.

3. Evaluating the flow characteristics in the individual ECs through fuzzy logic and identifying them as CSTR, PFR or dead volumes.

4. Agglomerating adjacent ECs that exhibit similar flow characteristics into compartments.

5. Defining possible connections between compartments based on the physical geometry of the system.

6. Estimating flow rates between linked compartments through an optimization based on the step response function of the system.

The steps of the proposed method and the results derived from it will be presented in a model system derived from the applications library of a commercial CFD simulator (COMSOL Multiphysics Version 5.2a). The top and side views of the system are shown in Fig. 3 with its characteristic dimensions.

To complete the first step of the algorithm, the CFD model of the unit was developed. The system was the model of a wastewater treatment reactor which originally consisted of four baffles, two of which were removed. To model the system, the fluid within the unit was assumed to exhibit the same properties as water. To investigate the behaviour of the system, the steady-state velocity field of the unit as well as the step response function of the system to an inert tracer were necessary.

To obtain these characteristics, the mass, impulse and component mass balance equations of the system have been solved numerically. The impulse balance of the system was solved assuming incompressible turbulent flow
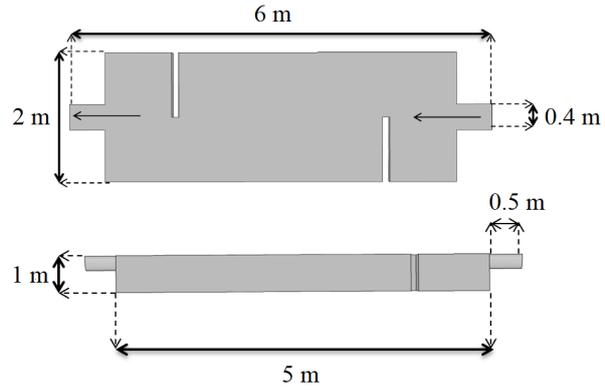
($Re_{\min} \approx 2 \cdot 10^5$ for the flow at the maximum diameter) within the unit using the $k - \epsilon$ turbulence model. The form of the Reynolds-averaged Navier-Stokes equations under these flow conditions is summarized in

$$\rho\left(\vec{v}\nabla\right)\vec{v} = \nabla\left[-p\vec{I} + \left(\mu + \mu_T\right)\left(\nabla\left(\vec{v}\right) + \left(\nabla\vec{v}\right)^T\right)\right] + F \tag{9}$$

$$\rho\nabla\vec{v} = 0 \tag{10}$$

$$\rho\left(\nabla\vec{v}\right)k = \nabla\left[\left(\mu + \frac{\mu_T}{\sigma_k}\right)\nabla k\right] + P_k - \rho\epsilon \tag{11}$$

$$\rho\left(\nabla\vec{v}\right)\epsilon = \nabla\left[\left(\mu + \frac{\mu_T}{\sigma_\epsilon}\right)\nabla\epsilon\right] + C_{\epsilon,1}\frac{\epsilon}{k}P_k - C_{\epsilon,2}\frac{\epsilon^2}{k} \tag{12}$$

$$\mu_T = \rho C_\mu \frac{k^2}{\epsilon} \tag{13}$$

$$P_k = \mu_T\left[\nabla\vec{v}\left(\nabla\vec{v} + \nabla\vec{v}^T\right)^{-1}\right] \tag{14}$$

The component mass balance and the step response function were calculated assuming that the velocity field within the unit had reached its steady state. No reactions were considered within the system and the change in the tracer concentration was attributed to convection and diffusion. The tracer was assumed to have the same physical properties as the water within the tank. Under these assumptions, the component mass balance within the unit can be expressed using the partial differential equation

$$\frac{\partial c}{\partial t} = -\vec{v}\nabla c - \nabla\left(-D\nabla c\right) \tag{15}$$

To solve the balance equations within the unit, a computational mesh was refined by taking into account both the computation time and accuracy of calculations, which was checked by calculating the relative mass balance error within the unit using

$$E_{\text{mass}} = \frac{\left(\sum_{i=0}^{n_{\text{sim}}} A_{\text{in}}\vec{v}_{\text{in},i}\rho_{\text{in}} - \sum_{i=0}^{n_{\text{sim}}} A_{\text{out}}\vec{v}_{\text{out},i}\rho_{\text{out}}\right)^2}{\max\left(\sum_{i=0}^{n_{\text{sim}}} A_{\text{in}}\vec{v}_{\text{in},i}\rho_{\text{in}}; \sum_{i=0}^{n_{\text{sim}}} A_{\text{out}}\vec{v}_{\text{out},i}\rho_{\text{out}}\right)} \tag{16}$$

Table 1: Parameters of the mesh

Cell type	Tetrahedral
Minimal size (mm)	23.1
Maximal size (mm)	77.4
E_{mass}	$5 \cdot 10^{-4}$

Table 2: Operating conditions of the unit

$T(°\text{C})$	20
Component	Water
$v_{\text{in}}\left(\text{m s}^{-1}\right)$	0.1
$p_{\text{out}}(\text{bar})$	1
$c_{\text{in}}\left(\text{mol m}^{-3}\right)$	2

The mesh was defined with a construction that allows for both an effectively short computation time and adequate degree of accuracy. The parameters of the mesh are shown in Table 1.

The steady-state velocity field and step response of the unit were calculated under the operating conditions listed in Table 2. After testing different numerical methods, it was found that in the case of the steady-state impulse balance calculations, the Generalized Minimal Residuals iterative algorithm was optimal for computations both with regard to computation time and computational accuracy. In the case of the component mass balance, by taking the same aspects into consideration, the conjugate gradient iterative method was utilized.

After solving the impulse balance of the unit under the previously displayed conditions, the streamlined steady-state velocity profile could be observed in Fig. 4. The streamlines were defined using the internal option of COMSOL, while their densities were set proportional to the local velocity field within an observed volume.

The figure shows that after the fluid enters the unit (on the left-hand side), various zones with different flow characteristics emerge due to the baffles. The flow entering the unit is at first unidirectional with PFR tendencies. The first baffle breaks this flow profile and creates a secluded area within which a well-mixed, circular, vortex-like flow could be observed. On the opposite side of the first baffle, a volume where small-scale mixing occurred could be observed. By following the main path of the flow, the area between the two baffles was a mixed regime with the presence of a significant dead volume. The area obstructed by the second baffle also showed dead volume tendencies. The flow above the second baffle assumed a tendency that was on the verge of being mixed and PFR-like with a clear PFR tendency dominating near the outlet. During the proposed method, the velocity field was converted into a vector field which is shown in Fig. 5. The vectors on the chart represent the tangents of the velocity streamlines within the unit. Their density, length and direction, as with the streamlines, was set to be proportional to the local magnitude and direction of the velocity
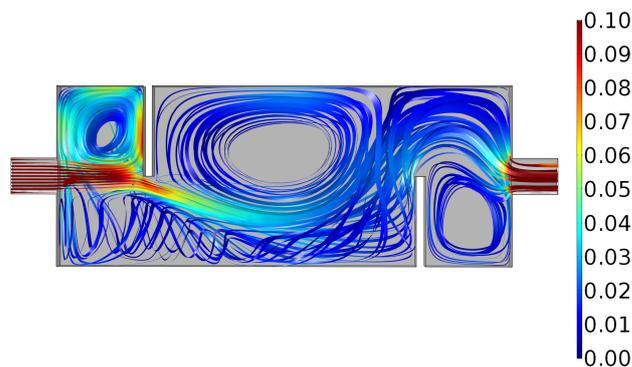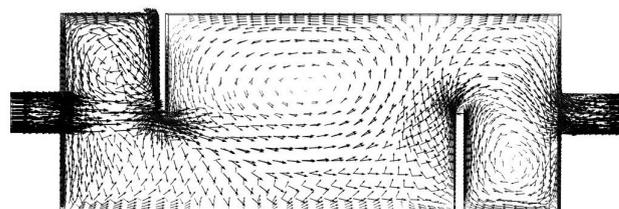

Figure 4: Steady-state streamlines within the unit $\left(\text{m s}^{-1}\right)$

Figure 5: Steady-state velocity vector field within the unit

field within the reactor.

The step function used to test the system and the response function of the unit are shown in Fig. 6. During the solution of the component mass balance equation diffusion coefficient $D = 10^{-4}\ \text{m}^2\ \text{s}^{-1}$ was used.

Based on the step response, the unit exhibited behaviour similar to a mixed CSTR unit combined with a PFR unit. The time delay between the step and response functions (0.05 h) is indicative of PFR behaviour within the system, while the curvature of the response shows CSTR tendencies. Irregularities and jumps within the response, as can be seen at approximately 0.08 h, might be due to internal circulation flows and backflow within the unit.
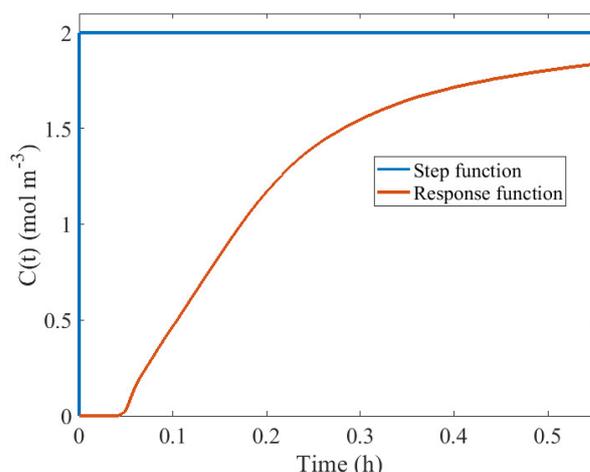

Figure 6: Step and response functions of the unit

3. Modelling results

Once the CFD results had been acquired, the results were processed using MATLAB R2020b. The geometry of the system was partitioned into EC units. Over the course of this investigation, the total volume of the tank reactor (4.96 m^3) was broken down into 25 rectangular ECs of equal volumes (0.2 m^3) and sizes in accordance with Step 2 of the proposed algorithm (Fig. 2). Investigations using different numbers of ECs were conducted and it was concluded that the most accurate results can be achieved by using a number of ECs where the size of the individual cells is approximately the same as that of the cells used in the CFD simulation. After the partitioning, the vectors characterizing the velocity field displayed in Fig.5 were observed within the ECs. Since the magnitude, direction and density of the vectors are all based on the attributes of the local velocity field within the tank, key observations can be derived from them about the local flow characteristics within the unit. The following general rules were established:

1. In areas where the flow exhibits PFR characteristics, the velocity vectors are mostly unidirectional with little variance in their magnitude. The magnitude of the velocity vectors is relatively large compared to the magnitude of the inlet velocity.

2. In areas where circular flow is present that exhibits CSTR characteristics, the variance with regard to the directional components of the velocity vectors is higher. The magnitude of the velocity vectors in these areas vary but are comparable to the magnitude of the inlet velocity.

3. In areas where little to no flow is present, also referred to as dead volumes, the magnitude of the velocity vectors is negligible compared to that of the inlet velocity.

Based on these rules, the flow characteristics within a certain volume of the system can be approximately categorized. Therefore, the velocity vectors within the individual ECs were observed and classified based on their flow characteristics. To achieve this, three measures were introduced. The average velocity within a cell (v_1), the variance in the directional components of the velocity vectors within a cell (v_2) and the skewness (v_3) with regard to the distribution of velocity vectors within a cell. To calculate the mean velocity within the cells, the total of the velocity vectors was calculated and averaged. Then the magnitude of the averaged vector was computed from

$$v_1 = \left| \frac{\sum_{i=1}^{n_{\vec{v}}} \vec{v}_i(x,y,z)}{n_{\vec{v}}} \right| \qquad (17)$$

The variance and skewness were calculated using similar approaches. Since the coordinates determine the directions, ECs that exhibit PFR behaviour where vectors are unidirectional will have a high mean velocity,

while those that contain circular flow with varying vector directions will generally have lower mean velocities. However, the variance will increase within ECs that exhibit strong circular flow with velocity vectors pointing in various directions. Dead volumes have exceptionally low mean velocities. In the case of all three variables, their numerical values were evaluated in the individual ECs. To properly compare these parameters, the numerical value of each variable within an EC was compared to its maximum value within the investigated system from

$$u_i = \frac{v_i}{\max(v_i)} \qquad (18)$$

Therefore, for each variable, a quantity was produced that determines its fractional value compared to its maximum value within the unit. The values of the three parameters within the system can be seen in Fig. 7.

Fig. 7 shows that the indicated variables can be used to differentiate between the flow behaviour of the unit based on the preliminary rules. The figure also explains the need for the inclusion of the skewness of the velocity distribution. Given that in the area adjacent to the first baffle, the flow velocity was rather high, the variance in the magnitude and direction of the velocity vectors in that area was exceptionally high. Compared to this value, most values of variance within the system were minuscule, except near the outlet where the circular flow met the plug-like flow exiting the reactor. The skewness parameter describes the asymmetry of a given distribution function. In areas with strong circular flow patterns, the directions of the velocity vectors are more evenly distributed on both sides of the mean velocity value, resulting in a more symmetric distribution of the velocity vector direction and lower skewness. However, in areas that exhibit plug flow-like behaviour, the distribution is generally more skewed due to the mostly unidirectional velocity vectors.

After obtaining the three quantities used to characterize the flow, their values within the individual ECs were evaluated in accordance with the previously defined rules to obtain the idealized EC structure mentioned in Step 3 of the algorithm (Fig. 2). During this process, the PFR, CSTR and dead volume characteristics of all the individual ECs were examined by converting the established rules using traditional fuzzy logic [23]. Fuzzy expert systems utilize empirical knowledge from subject-matter experts to characterize and analyze processes [24].

They contain a working memory, inputs, a collection of rules and outputs. The working memory is the culmination of objects O_i, where **O** denotes the set of all objects. Objects consist of a fuzzy vector ($\vec{V}_i$) where all members of the vector are fuzzy sets or fuzzy numbers $\left(\vec{V}_{i,1}, \vec{V}_{i,2}, ..., \vec{V}_{i,n}\right)$. These fuzzy sets describe qualitative information about the objects, which in this case were each of the individual ECs. The inputs of the fuzzy expert system were the variables u_i, which were used to describe the flow characteristics within an individual EC.
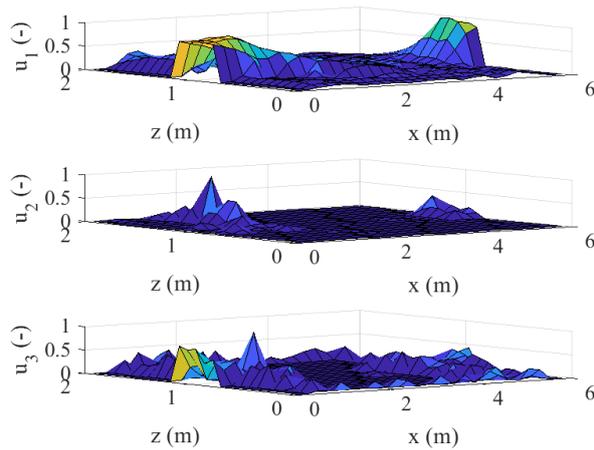
Figure 7: Characteristic velocity variables within the system

f_{input}	a	b	f_{output}	a	b
$\text{Input}_{1,\text{low}}$	0.01	0.05	$\text{Output}_{\text{low}}$	0.02	0.4
$\text{Input}_{2,\text{low}}$	0.01	0.05			
$\text{Input}_{3,\text{low}}$	0.02	0.1			
$\text{Input}_{1,\text{high}}$	0.03	0.2	$\text{Output}_{\text{high}}$	0.02	0.6
$\text{Input}_{2,\text{high}}$	0.02	0.2			
$\text{Input}_{3,\text{high}}$	0.03	0.3			

puts of our object, which in our case meant the characteristics of the PFR, CSTR and dead volume of individual ECs [25]. Given that the linguistic variables $P_2 = \{\text{high}, \text{low}\}$ are represented by membership function $f_{\text{output},j}$ for each characteristic, the membership functions are monotonic and sigmoidal like the input membership functions as shown in

$$f_{\text{output},j}\left(f_{\text{input},j}\left(R\right)\right) = \frac{1}{1 + e^{-\left(\frac{f_{\text{input},j}(R) - b}{a}\right)}} \quad (21)$$

$$f_{\text{output},j}\left(f_{\text{input},j}\left(R\right)\right) = 1 - \frac{1}{1 + e^{-\left(\frac{f_{\text{input},j}(R) - b}{a}\right)}} \quad (22)$$

In this equation, $f_{\text{input},j}\left(R\right)$ represents the degree of belief in R for each subset. After defuzzification, the expert system returned crisp values that determined the PFR, CSTR and dead-volume characteristics within each EC.

The input and output membership functions can be observed in Fig. 8 and Fig. 9, respectively. Since the output membership functions in the case of all three outputs were identical, they have only been displayed once. The first index within the legend represents the input which the membership function belongs to, while the second one corresponds to the linguistic subsets ($1 \rightarrow$ Low, $2 \rightarrow$ High) [25].
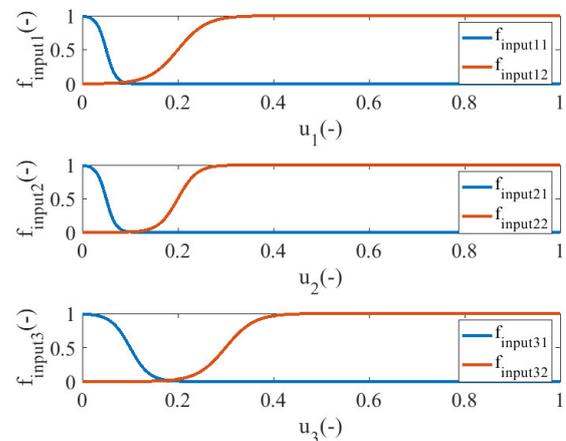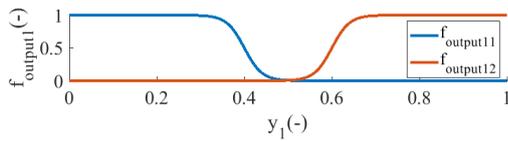
A set of rules ($\mathbf{R}$) was established, which in this case pertains to the flow characteristics within the unit. A fuzzy subset of $\mathbf{R}$ denoted as R with the membership function f_{input} provides $f_{\text{input}}\left(R\right)$, which symbolizes the level of belief in R for each subset [25]. The degree of belief is a number between 0 and 1 with 0 and 1 denoting disbelief and wholehearted belief in the proposed statement, respectively. The rules are obtained in the form of "IF-THEN" rules with logical operators like "OR", "AND", etc. In the case of our system, the real number inputs u_i were converted into fuzzy numbers using monotonic sigmoidal membership functions, their general form is presented

$$f_{\text{input},j}\left(u_i\right) = \frac{1}{1 + e^{-\left(\frac{u_i - b}{a}\right)}} \quad (19)$$

$$f_{\text{input},j}\left(u_i\right) = 1 - \frac{1}{1 + e^{-\left(\frac{u_i - b}{a}\right)}} \quad (20)$$

for fuzzy numbers with index j generated from input i.

The first equation is used for membership functions which are increasing in tendency, while the latter is for membership functions which are decreasing. Parameters a and b define the spread of the fuzzy numbers either side of their central value and the central values of the fuzzy numbers, respectively. The fuzzy numbers obtained through the membership functions represent the linguistic variables P belonging to the set of basic linguistic variables $\mathbf{P}$. In our case, these linguistic variables consisted of the set $P_1 = \{\text{high}, \text{low}\}$ which was used to describe whether the mean value, variance and skewness of velocity vectors within an EC were high or low compared to the maximum value of the respective attribute within the system.

The output actions define a fuzzy set of conclusions. The set $\mathbf{G}$ represents the set of possible conclusions of the expert system. A subset of this set represents the out-

Figure 8: Input membership functions

Figure 9: Output membership functions

Figure 10: Flow characteristics of the different EC units

Figure 11: Characteristic flow behaviour within the unit

The parameters a and b of the membership functions are displayed in Table 3. Using the expert system, the flow characteristics of the ECs were characterized. The PFR (y_1), CSTR (y_2) and dead volume (y_3) characteristics within the ECs are shown in Fig. 10.

To achieve a flow map that describes the behaviour of the velocity field within each unit, the three characteristics were unified. The weighted average of the three output characteristics was computed and a function defined by $\text{Im}(L) \in [0, 2]$ where $L(y_1, y_2) = 0$ corresponds to the pure dead-volume behaviour, $L(y_1, y_2) = 1$ denotes the pure CSTR behaviour and $L(y_1, y_2) = 2$ stands for the pure PFR behaviour. $L(y_1, y_2)$ in each individual EC was calculated from

$$L(y_1, y_2) = \frac{2 \cdot y_1 + y_2}{3}. \tag{23}$$

Since the weighting factor of the dead-volume behaviour is 0, it was not included in the equation. The results of the process are shown in Fig. 11.

The correlation between Fig. 4 and Fig. 11 is apparent. The model predicts clear PFR behaviour within the flow near the inlet and outlet which can also be observed from Fig. 4. The area beyond the inlet and bounded by the first baffle showed CSTR tendencies with some PFR behaviour near the walls. A dead volume could be observed at the centre of the stirred volume which was similar to in the middle of the vortex seen in Fig. 4. The area opposite the volume bounded by the first baffle exhibited dead-volume tendencies with a low velocity and occasional circular flow patterns, indicative of slight CSTR behaviour. The area between the two baffles showed mixed, CSTR
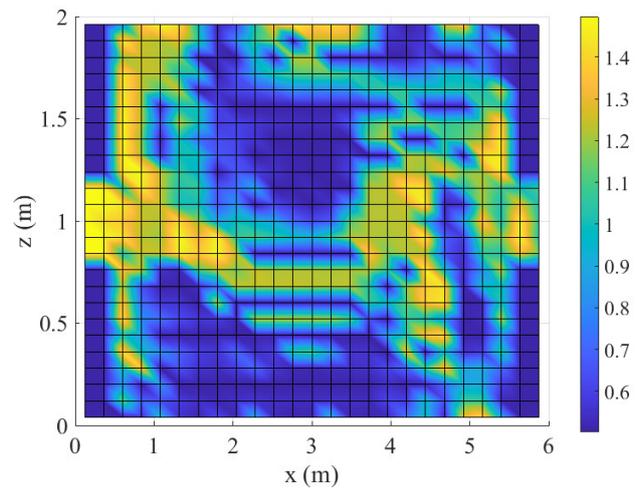
tendencies with a noticeable dead volume present in the middle of the vortex, as can be seen in Fig. 4. The area bounded by the second baffle mainly exhibited dead-volume tendencies with a medium degree of circulation. The volume opposite it consisted of a mixed flow. The flow near the outlet, as noted before, exhibited PFR tendencies.

The figure presents an approximately accurate description of the flow characteristics when compared to the streamlines obtained from the CFD methods. However, due to fuzzification and the weighted average calculations to ascertain the flow tendencies within the unit, individual flow characteristics of the ECs are unclear. To clarify these values, limits were introduced to clearly categorize the flow behaviour of all ECs, which were calculated according to Eq. 24. The limits for this clustering process were determined by observing flow patterns obtained using various limits. It was found that these limits facilitated the most efficient clustering and resulted in a cluster which strongly resembled the flow characteristics observed in the CFD results.

$$L_{\text{disc}}(y_1, y_2) = \begin{cases} 2, & \text{if } L(y_1, y_2) > 1.25 \\ 1, & \text{if } 1.25 \geq L(y_1, y_2) \geq 0.75 \\ 0, & \text{otherwise} \end{cases} \tag{24}$$

This equation shows that the values of the corrected function are discrete, that is, the ECs are categorized into three idealized classes of flow depending on which pure idealized class their behaviour most resembles.

The plot displaying the flowchart with ECs containing the corrected values is shown in Fig. 12. The figure also indicates the locations of the inlet, outlet and baffles. Since a rectangular mesh was used to divide the system into ECs, the locations that were not part of the geometry and the baffles were also included in the calculations. Since no flow was present at these locations, they have been categorized as clear dead volumes, which can be seen by comparing Fig. 11 and Fig. 12. For the sake of clarity, these areas were removed from the corrected fig-
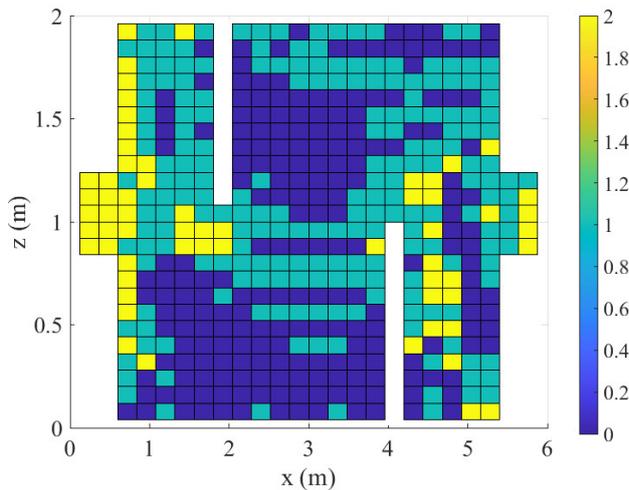
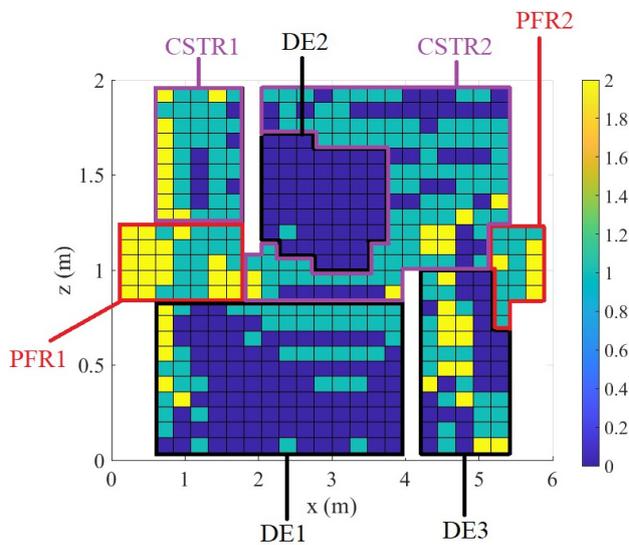*Figure 12:* Recognized flow behaviour within the ECs

Figure 14: Estimated CM structure

in Fig. 14.

The constructed CM structure contains the previously identified compartments seen in Fig. 13. Connections between the different compartments were established through the use of theoretical mixer and divider units. During the evaluation of the possible connections between compartments, the geometry of the system was taken into account as well, e.g., the placement of baffles, positions of adjacent compartments, etc. The models of the individual compartments, mixers and dividers were formulated according to the equations provided in the Introduction.

After obtaining the CM structure, Step 6 of the identification algorithm (Fig. 2) involved estimating the flow rates between compartments. In the case of the structure shown in Fig.14, the flow rates between compartments were defined through the division rates (α_i) of the D units. These values describe the ratio of mass flow entering the ith adjacent compartment compared to the total mass flow leaving the jth reference compartment.

For the identification of the α_i values optimization was conducted in MATLAB R2020b through the Simulink interface. The constructed CM structure was defined in Simulink along with its connections and individual compartments. The system of differential and algebraic equations describing the behaviour of the system were solved numerically using Rosenbrock methods for solving stiff differential equations. The step response of the CM was observed as a function of the α_i values and an objective function was formulated to estimate the α values displayed in

$$E_C(\vec{\alpha}) = \frac{\sum_{i=1}^{n_{\text{sim}}} (C_{\text{CFD},i} - C_{\text{CM},i})^2}{\max \left(\sum_{i=1}^{n_{\text{sim}}} C_{\text{CFD},i}; \sum_{i=1}^{n_{\text{sim}}} C_{\text{CM},i} \right)} \quad (25)$$

The objective function estimates the squared difference between the step response function acquired through CFD methods and the CM over the observed discrete-time horizon relative to the maxima of the sum of the two functions. To minimize the objective function with regard to the α_i values, the interior-point method was utilized. The step response function of the system obtained through CFD methods and the optimized CM can be seen in Fig. 15.

The resulting figure shows that the step response of the CM after optimization fits the step response of the system obtained through CFD methods well with an aver-

Figure 13: Defined compartments within the system

ure.

Fig. 12 shows a clear map indicating the idealized flow behaviour of individual ECs within the unit. Once agglomerated, adjacent ECs that exhibited similar characteristics could be grouped to form compartments of idealized behaviour in accordance with Step 4 of the proposed algorithm (Fig. 2). The agglomeration of the ECs within this paper was conducted manually. The resulting set of compartments is shown in Fig. 13.

The unit was partitioned into seven compartments, namely two PFR regions, two CSTR regions and three dead volumes. The volumes of the compartments were evaluated by totalling the volumes of the individual ECs which were assigned as part of the compartment in question. It can be seen that the resulting structure of the compartment strongly resembles the results obtained through CFD methods seen in Fig. 4. Based on the defined structure, a CM consisting of the compartments that exhibited idealized behaviour was constructed to complete Step 5 of the algorithm (Fig. 2). The CM structure is displayed
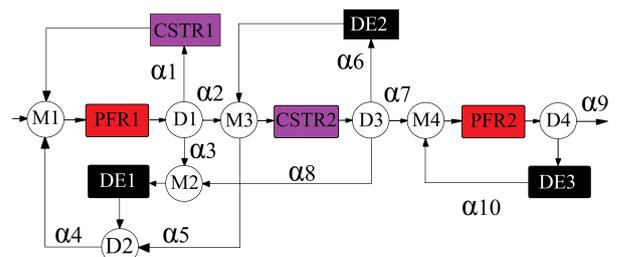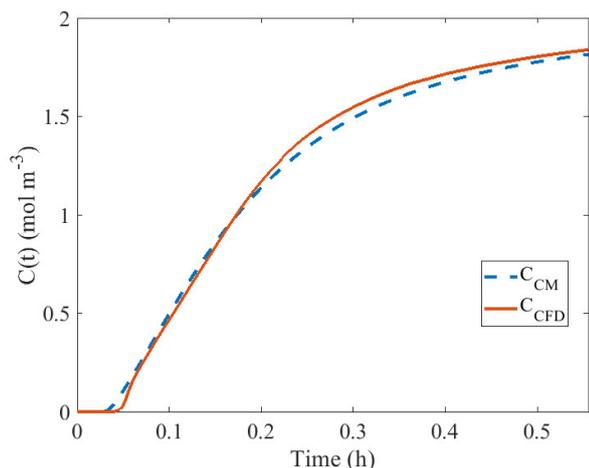
Figure 15: Step response of the system with CFD methods and CM

Table 4: Optimized α parameters

α_1	0.7	α_6	0
α_2	0.25	α_7	1
α_3	0.05	α_8	0
α_4	0.8	α_9	0.95
α_5	0.2	α_{10}	0.05

age fit of 92%. The optimized model parameters rounded up to two decimal places are presented in Table 4.

The estimated parameters in Table 4 are also in line with the general CM structure. It should be noted that the optimized α parameters of flows entering dead volumes $(\alpha_3, \alpha_6, \alpha_8, \alpha_{10})$ were exceptionally low, moreover, sometimes not even detectable when rounded up to two decimal places. This indicates that the step response of the system is most similar to the response obtained through CFD methods if almost no flow enters the listed volumes, indicating that they are truly dead volumes from a flow perspective.

4.   Conclusions

Over the course of this study, the process of creating CMs from results obtained through CFD simulations has been presented. The proposed algorithm consisted of two phases. In the first phase, the CM structure of the object was defined. For these investigations, the steady-state velocity field of the system was utilized. The system was partitioned into ECs in which the local velocity field was investigated. By analyzing the direction and magnitude of local velocity vectors within the ECs using fuzzy logic, a set of rules was established to categorize the behaviour of individual ECs and correlate these ECs to units exhibiting idealized behaviour.

To characterize the flow behaviour, three measures were defined based on the distribution of the magnitude and direction of velocity vectors within individual ECs. Three types of ECs were considered, ones in which the

flow is well-mixed and circular, similar to a CSTR, ones in which the flow is unidirectional and resembles that found in a PFR, and ones in which little to no flow is present, that is, dead-volume behaviour is exhibited.

After the qualitative analysis using fuzzy logic, each EC was assigned an idealized trait. Adjacent ECs that exhibited similar traits were clustered into compartments showing idealized behaviour. After determining the idealized compartments, the flow rates between these compartments were estimated through optimization by calculating the flow rates for which the sum of the squared difference between the step response of the CFD model and CM was minimal. The estimated parameters and the observed compartments showed a strong correlation with the flow patterns which could be observed within the CFD results. The method is capable of effectively analyzing CFD results and converting them into CMs in a structured framework while also allowing the addition of empirical knowledge using fuzzy logic. The proposed method could serve as a basis for the estimation of robust and computationally inexpensive a priori models to describe equipment utilized within the chemical industry. The physical correlation between the model compartments and the actual flow regimes within the unit could provide opportunities to develop a priori model-based diagnostic strategies for online fault detection among many other applications.

5.   Acknowledgements

Nomenclature

Fuzzy sets

$\vec{\mathbb{V}}$	Set of fuzzy vectors/numbers
$\mathbb{G}$	Set of possible fuzzy conclusions
$\mathbb{O}$	Set of fuzzy objects
$\mathbb{P}$	Set of linguistic fuzzy variables
$\mathbb{R}$	Set of fuzzy rules

Greek letters

α	Distribution ratio $[-]$
ϵ	Turbulent dissipation rate $[\mathrm{m^2\,s^{-3}}]$
μ	Kinematic viscosity $[\mathrm{m^2\,s^{-1}}]$
μ_T	Turbulent viscosity $[\mathrm{kg\,m^{-1}\,s^{-1}}]$
ρ	Density $[\mathrm{kg\,m^{-3}}]$
σ_ϵ	Constant for change in ϵ $[-]$
σ_k	Constant for change in k $[-]$
τ	Residence time $[\mathrm{h}]$

Latin letters

$\vec{F}$	Force $[\mathrm{N}]$

$\vec{I}$ Impulse $[\mathrm{kg\,m\,s^{-1}}]$

$\vec{v}$ Velocity $[\mathrm{m}\,s^{-1}]$

A Surface $[\mathrm{m^2}]$

a Fuzzy membership function central value $[-]$

b Fuzzy membership function variance $[-]$

C Step response function $[\mathrm{mol\,m^{-3}}]$

c Species concentration $[\mathrm{mol\,m^{-3}}]$

$C_{\epsilon,1}$ Dissipation rate calculation constant 1 $[-]$

$C_{\epsilon,2}$ Dissipation rate calculation constant 2 $[-]$

C_μ Turbulent viscosity calculation constant $[-]$

D Diffusion rate $[\mathrm{m^2\,s^{-1}}]$

E Error function $[-]$

F Flow rate $[\mathrm{m^3\,s^{-1}}]$

f Fuzzy membership function

H Heaviside (Step) function $[\mathrm{mol\,m^{-3}}]$

i, j Numeric indices $[-]$

k Turbulent kinetic energy $[\mathrm{m^2\,s^{-2}}]$

L Continuous flow characteristic function $[-]$

L_{disc} Discretized flow characteristic function $[-]$

n Absolute number $[-]$

p Pressure $[\mathrm{Pa}]$

P_k Turbulent kinetic energy due to mean velocity variations $[\mathrm{kg\,m^{-1}\,s^{-1}}]$

T Temperature $[^\circ\mathrm{C}]$

t Time $[\mathrm{h}]$

u Fuzzy expert system input variable $[-]$

V Volume $[\mathrm{m^3}]$

x Longitudinal coordinate $[\mathrm{m}]$

y Fuzzy expert system output variable $[-]$

z Lateral coordinate $[\mathrm{m}]$

REFERENCES

[1] Venkatasubramanian, V.; Rengaswamy, R.; Yin, K.; Kavuri, S.N.: A review of process fault detection and diagnosis: Part I: Quantitative model-based methods., *Comput. Chem. Eng.*, 2003, **27**(3), 293–311 DOI: 10.1016/S0098-1354(02)00160-6

[2] Broughton, E.: The Bhopal disaster and its aftermath: a review., *Environ. Health*, 2005, **4**(1), 1–6 DOI: 10.1186/1476-069X-4-6

[3] Isimite, J.; Rubini, P.: A dynamic HAZOP case study using the Texas City refinery explosion., *J. Loss. Prev. Process. Ind.*, 2016, **40**(1), 496–501 DOI: 10.1016/j.jlp.2016.01.025

[4] J., G.: Fault detection and diagnosis in engineering systems. (CRC press), 1998. ISBN: 9780367400439

[5] Venkatasubramanian, V.; Rengaswamy, R.; Kavuri, S.N.: A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies., *Comput. Chem. Eng.*, 2003, **27**(3), 313–326 DOI: 10.1016/S0098-1354(02)00161-8

[6] Venkatasubramanian, V.; Rengaswamy, R.; Kavuri, S.N.; Yin, K.: A review of process fault detection and diagnosis: Part III: Process history based methods., *Comput. Chem. Eng.*, 2003, **27**(3), 327–346 DOI: 10.1016/S0098-1354(02)00162-X

[7] Solin, P.; Segeth, K.; Dolezel, I.: Higher-order finite element methods. (CRC press), 2003. ISBN: 9780429205279

[8] Bezzo, F.; Macchietto, S.; Pantelides, C.C.: A general methodology for hybrid multizonal/CFD models: Part I. Theoretical framework., *Comput. Chem. Eng.*, 2004, **28**(4), 501–511 DOI: 10.1016/j.compchemeng.2003.08.004

[9] Savic, R.M.; Jonker, D.M.; Kerbusch, T.; Karlsson, M.O.: Implementation of a transit compartment model for describing drug absorption in pharmacokinetic studies., *J. Pharmacokinet. Pharmacodyn.*, 2007, **34**(5), 711–726 DOI: 10.1007/s10928-007-9066-0

[10] Pigou, M.; Morchain, J.: Investigating the interactions between physical and biological heterogeneities in bioreactors using compartment, population balance and metabolic models., *Chem. Eng. Sci.*, 2015, **126**(1), 267–282 DOI: 10.1016/j.ces.2014.11.035

[11] Spann, R.; Gernaey, K.V.; Sin, G.: A compartment model for risk-based monitoring of lactic acid bacteria cultivations., *Biochem. Eng. J.*, 2019, **151**(1), 2–11 DOI: 10.1016/j.bej.2019.107293

[12] Kaur, G.; Singh, M.; Matsoukas, T.; Kumar, J.and De Beer, T.; Nopens, I.: Two-compartment modeling and dynamics of top-sprayed fluidized bed granulator., *Appl. Math. Model.*, 2019, **68**(1), 267–280 DOI: 10.1016/j.apm.2018.11.028

[13] Egedy, A.; Varga, T.; Chován, T.: Compartment model structure identification with qualitative methods for a stirred vessel., *Math. Comput. Model. Dyn. Syst.*, 2013, **19**(2), 115–132 DOI: 10.1080/13873954.2012.700939

[14] Levenspiel, O.: Chemical reaction engineering. (John Wiley & Sons), 1999. ISBN: 978-0-471-25424-9

[15] Claudel, S.; Fonteix, C.; Leclerc, J.P.; Lintz, H.G.: Application of the possibility theory to the compartment modelling of flow pattern in industrial processes., *Chem. Eng. Sci.*, 2003, **58**(17), 4005–4016 DOI: 10.1016/S0009-2509(03)00269-0

[16] Delafosse, A.; Collignon, M.L.; Calvo, S.; Delvigne, F.; Crine, M.; Thonart, P.; Toye, D.: CFD-based compartment model for description of mixing in bioreactors., *Chem. Eng. Sci.*, 2014, **106**(1), 76–85 DOI: 10.1016/j.ces.2013.11.033

[17] Fogarasi, S.; Egedy, A.; Imre-Lucaci, F.; Varga, T.; Chován, T.: Hybrid CFD-compartment approach for modelling and optimisation of a leaching reactor., *Comput. Aided Chem. Eng.*, 2014, **33**(1), 1255–1260 DOI: 10.1016/B978-0-444-63455-9.50044-1

[18] Nauha, E.K.; Kálal, Z.; Ali, J.M.; Alopaeus, V.: Compartmental modeling of large stirred

tank bioreactors with high gas volume fractions., *Chem. Eng. J.*, 2018, **334**(1), 2319–2334 DOI: 10.1016/j.cej.2017.11.182

[19] Weber, B.; von Campenhausen, M.; Maßmann, T.; Bednarz, A.; Jupke, A.: CFD based compartment-model for a multiphase loop-reactor., *Chem. Eng. Sci.: X*, 2019, **2**, 100010 DOI: 10.1016/j.cesx.2019.100010

[20] Bezzo, F.; Macchietto, S.: A general methodology for hybrid multizonal/CFD models: Part II. Automatic zoning., *Comput. Chem. Eng.*, 2004, **28**(4), 513–525 DOI: 10.1016/j.compchemeng.2003.08.010

[21] Tajsoleiman, T.; Spann, R.; Bach, C.; Gernaey, K.V.; Huusom, J.K.; Krühne, U.: A CFD based automatic method for compartment model development., *Comput. Chem. Eng.*, 2019, **123**(1), 236–245

DOI: 10.1016/j.compchemeng.2018.12.015

[22] Nørregaard, A.; Bach, C.; Krühne, U.; Borgbjerg, U.; Gernaey, K.V.: Hypothesis-driven compartment model for stirred bioreactors utilizing computational fluid dynamics and multiple pH sensors., *Chem. Eng. J.*, 2019, **356**(1), 161–169 DOI: 10.1016/j.cej.2018.08.191

[23] Zadeh, L.: Fuzzy logic., *Computer*, 1988, **21**(4), 83–93 DOI: 10.1109/2.53

[24] Hadjimichael, M.: A fuzzy expert system for aviation risk assessment., *Expert Syst. Appl.*, 2009, **36**(3), 6512–6519 DOI: 10.1016/j.eswa.2008.07.081

[25] Buckley, J.J.; Siler, W.; Tucker, D.: A fuzzy expert system., *Fuzzy Sets Syst.*, 1986, **20**(1), 1–16 DOI: 10.1016/S0165-0114(86)80027-6