SI-M/O: Swarm Intelligence-based Modelling and Optimization of complex reaction processes

Min Wu[1], Ulderico Di Caprio[1], Furkan Elmaz[2], Florence Vermeire[3], Bert Metten[4], Olivier Van Der Ha[4], Dries De Clercq[4], Siegfried Mercelis[2], Peter Hellinckx[2], Leen Braeken[1], M. Enis Leblebici[1,*]

[1] Center for Industrial Process Technology, KU Leuven, Diepenbeek, Belgium

[2] Faculty of Applied Engineering, University of Antwerp, Antwerp, Belgium

[3] Chemical Reactor Engineering and Safety, KU Leuven, Leuven, Belgium

[4] Ajinomoto Bio Pharma Services, Wetteren, Belgium

*Corresponding author (muminenis.leblebici@kuleuven.be)

Abstract

A Swarm Intelligence-based Modelling and Optimization (SI-M/O) algorithm is proposed in this paper, consisting of two parts: SI-M and SI-O. Processes with unknown reactions and limited data are challenging to model with high accuracy. SI-M uses a hybrid modelling approach via swarm intelligence algorithms and simple polynomial regression terms. It achieves high accuracy and wide applicability with a reasonable computational time. Based on the developed models, SI-O for a production plant is introduced. The productivity in the preliminary plant test was increased by 5.3% in one minute. SI-O not only maximizes the productivity, but also ensures the continuous plant operations. In addition, a graphical user interface of SI-M/O is designed for an easy usage.

Keywords: swarm intelligence, hybrid modelling, process optimization, continuous flow reactor, graphical user interface

1. Introduction

A novel method of Swarm Intelligence-based Modelling and Optimization (SI-M/O) for complex reaction processes is proposed in this paper. SI-M/O consists of two major parts: SI-M for process modelling and SI-O for process optimization. Process modelling and simulation are essential for digitization in modern chemical industry. Generally, the approaches of modeling and simulation are divided into three categories: first-principle, data-driven, and hybrid (Zendehboudi et al., 2018). All of the three approaches have got decent performances in modelling reaction processes (Babanezhad et al., 2020; Rojnuckarin et al., 1993; Xie et al., 2018). First-principle modelling is based on chemical knowledge and provides high extrapolation capabilities. However, it requires the model designer to have a deep knowledge about the system taken into account (Pantelides and Renfro, 2013). This modelling approach is challenging, or even impossible, if the system is partially unknown. For example, there are unknown intermediates, reactions or interactions in a reaction process. The data-driven modelling derives information directly from data, but it lacks physical consistency, and has poor capabilities of generalization and extrapolation. The ideal dataset utilised in the data-driven modelling should have high '4V' characteristics –Volume, Velocity, Veracity, and Variety (Anagnostopoulos et al., 2016), which is possible to acquire in simulations but is rare in laboratory or plant experiments. The data volume refers to the involved data quantity. A lack of data volume leads to unreliable derived information for modelling, but massive data volume results in data preproccessing problems. Data velocity shows the speed of getting and feeding new data. Data veracity defines the degree of involved data accuracy and precision, and data variety is the diversity of collected data in the problem space. In chemical engineering, it is common that the data from laboratories and plants is restricted to a specific range or a certain point because of practical and safety issues, which results in low data variety. Therefore, to overcome data and knowledge problems in complex reaction processes, a hybrid approach is preferred, which is a combination and takes both the advantages of first-principle and data-driven methods.

SI-M proposed in this paper is a hybrid approach, combining experimental data with low 4Vs and basic first principles that express mass or energy balances to simulate complex reaction processes. In the open literature Hybrid modelling has been widely applied for chemical processes simulations (Sansana et al., 2021; von Stosch et al., 2014). Psichogios et al. developed a hybrid model for a fed-batch bioreactor with Artificial Neural Networks (ANN) (Psichogios and Ungar, 1992). In their research, they showed how the hybrid modelling outstand pure data-driven techniques. Following their research, the ANN-based hybrid modeling framework was employed by numerous authors, such as for Baker's yeast production (Feyo De Azevedo et al., 1997), Fischer-Tropsch reaction (Shiva et al., 2013), and esterification reaction (Ammar et al., 2021). However, these ANN-based modelling approaches still required a large amount of data with high 4Vs or careful consideration when applied outside the training domain. Besides ANN, other machine learning techniques were also investigated in hybrid modelling. For example, partial least square-based hybrid modelling was applied to study the concept of the hybridization degree for cell culture process modelling (Narayanan et al., 2022), and a hybrid technique integrating principal component analysis was introduced to detect faults in heating, ventilation, and air conditioning systems (Hassanpour et al., 2020). To the best of our knowledge, there are few studies for complex reaction processes to overcome both difficulties simultaneously: a) unknown reactions and intermediates; b) a low amount of available data. In many cases, pure first-principle modelling is impossible because of the lacking knowledge of reactions or reactors. A huge amount of time and effort is needed to research unknown reactions or intermediates and to study the first priciples. Due to the low availability of data 4Vs, advanced machine learning techniques (e.g. ANN) would be unsuitable and easily result in overfitting. Therefore, SI-M in this paper uses simple polynomial regression terms to reduce the number of parameters. This novel hybrid modelling approach integrated with Swarm Intelligence (SI) algorithms to enhance the extrapolation capabilities of models, to reduce the amount of required data, and to decrease the time of model development.

As a subfield of AI, SI inspired by biological swarm behaviours is gaining more attention. It is collective intelligence collaborating with relatively simple individuals or processing units (Chakraborty and Kar, 2017). SI algorithms are named after the inspiration sources, and have been applied in chemical engineering. For example, Partical Swarm Optimisation (PSO) showed its efficiency for constructing confidence regions in a polymerization kinetics problem (Schwaab et al., 2008), where the first-principle model describing the rates of ethylene consumption was given. In this paper, SI algorithms estimiated the parameters in not only first-principle models but also data-driven models. It is hypothsized that SI algorithms can solve the problems of high-dimensional parameters. PSO was also used in process optimization to obtain the optimal process operations (Fang et al., 2021). In this paper, the SI-based Optimisation (SI-O) approach for a continuous production plant is proposed. The aim of the proposed SI-O is to maximize the productivity and to ensure the plant operates continuously. It is hypothesized that SI algorithms can reach the relatively global optima instead of trap at local optima. Besides PSO, the applications of other SI algorithms in chemical engineering have also become increasingly common. For instance, an Ant Colony Optimisation (ACO)-extended methodology was successfully applied in wastewater treatment plants for process design and control (Schlueter et al., 2009). The authors find there are few Slime Mould Algorithm (SMA) applications in chemical engineering. In contrast, its applications in electronic engineering (Kumar et al., 2020) and industrial engineering (Abdel-Basset et al., 2021) have shown a decent performance. It is noted that in this paper three SI algorithms PSO (animal-based), ACO (insect-based) and SMA (bacteria-based) were investigated for both process modeling (SI-M) and process optimization (SI-O), avoiding the similarities within the same group of SI categories.

In this paper, SI-M/O (Swarm Intelligence-based Modelling and Optimization) of complex reaction processes is proposed. The system analyzed in this work shows high uncertainties where unknown reaction paths and intermediates exist. This makes impossible to create reliable and general first-principle models. Furthermore, the data used in this study for the model training did not respect the ‘4V’ characteristics. Therefore, SI-M (Swarm Intelligence-based Modelling) is applied to rapidly develop an accurate model with limited data. Following the model development, SI-O (Swarm Intelligence-based Optimization) is introduced to is to maximize the productivity and to ensure the plant operates continuously. A graphical user interface is designed to display the performance of SI-M/O.

2. Methodology

The data utilised in this paper are from experiments in chemical laboratories and production plants. Because the experiments are time-consuming and there are strict safety regulations, the obtained data has a low volume and a low variety. Also, the reaction process consists of unknown reactions and intermediates. However, the proposed SI-M (Swarm Intelligence-based Modelling) with simple polynomial regression terms can be used to overcome the challenges. Furthermore, SI-O is introduced to optimize the system.

2.1. Swarm intelligence

Swarm intelligence (SI) can be sorted into meta-heuristic search algorithms. Their search is not based on the gradient of the function but relies on the synergic work of several particles sharing information to reach the optima of the optimization function. One of the main advantages of SI techniques is that the cost function can also be discontinuous (Kennedy, 2006).The initial swarms are generated, starting within the assigned search space. The evaluation of each individual is then made to check whether it meets the requirements. Each unit in the population would be updated by the policies inspired by nature swarms until the criteria are satisfied. The new individuals would repeat the same evaluation and renewal process until they meet the requirements, and the final units can be seen as acceptable solutions.

In this paper, three SI algorithms are applied, including Slime Mould Algorithm (SMA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO). SMA is from the approaching food method of slime mold according to the odour in the air (Li et al., 2020). PSO is inspired by the bird swarm behaviour proposed in 1995 (Kennedy and Eberhart, 1995). The initial ACO algorithms are designed to solve discrete optimization problems and have been extended to continuous domains since 1995 (Blum, 2005). This paper uses the extended ACO algorithm through the probabilistic sampling of search spaces (Riadi, 2014).

2.2. Process modelling

In this paper, the reactions involved in the processes are complex, while the experimental data is limited. SI-M is presented to overcome these obstacles for achieving an accurate model from the limited data. The training structure of the SI-M is shown in Fig. 1a, including data, models, objective function, and swarm intelligence. The aim is to get the minimum value of the objective function (i.e., objective value). The objective values are calculated from the experimental outputs, the modelled outputs, and the parameters. At every optimization step, SI-M suggested another set of parameters within the pre-defined boundaries. The model is simulated using the proposed parameters. From the simulation results, the objective value is calculated and the loop continues until the stopping criteria are met. The optimal set of parameters is obtained when the training stops.

The hybrid models consist of reaction models and reactor models. Fig. 1b gives the elements and parameters in hybrid models, including reaction and reactor models. The involving residual terms are the simple polynomial regression, which is the black-box modelling part in SI-M. The corresponding parameters are the regression factors, and the details are shown in Equation 20, Equation 21, and Equation 22. The white-box models are derived from known reactions and measurable chemicals. The employed equations are based on Arrenhuis equations and steady-state plug-flow reactor models, where the incorporated parameters can be found in Equation 7 and Equation 16. The parameters in the white-box part are called white-box parameters, and the ones in the black-box part are black-box parameters. This work applies training approaches: 1) the white-box and black-box parameters are identified at the same time; 2) the white-box parameters are identified as first followed by another training to identify the parameters of the ML polynomial functions. The two training approaches are called one-time and step-by-step, respectively.

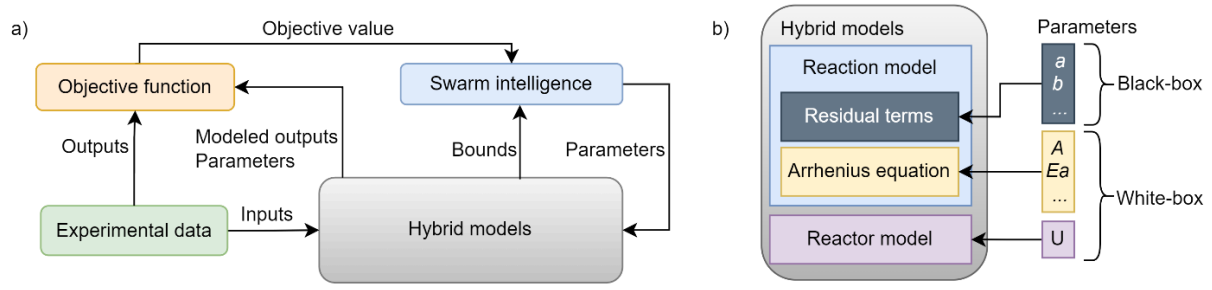

Fig. 1: a) The training structure of SI-M (Swarm Intelligence-based Modelling). It consists of four parts: data, models, objective function, and swarm intelligence algorithms. b) The elements and parameters in hybrid models, including reaction and reactor models. The involving residual terms play a role as the black-box modelling part, with parameters a, b, …(see Equation 20, Equation 21, Equation 22). The white-box models are incorporated with A (pre-exponential factor), Ea (activation energy), U (heat transfer coefficient) (see Equation 7, Equation 16).

2.2.1. First-principle models

A reaction process model can be divided into reaction kinetics and reactor models. There were three types of experiments conducted - Continuous Flow Reactor (CFR), Semi-Batch Reactor (SBR), and Reaction

Calorimetry (RC1) to acquire the data. Because the limiting information was obtained from CFR processes where only the reactor outlet can be measured, SBR and RC1 experiments were also conducted to investigate the reactions. RC1 experiments were conducted in SBR operation to determine the reaction heats. SBR experiments provide more information about side reactions at the long residence time. RC1 experiments give observed reaction heats of different reactions. Moreover, the CFR experiments focus on reactions at a short residence time and their performance at various temperatures. The reactions in the studied processes are complex, and some are unknown, as shown in Fig. 2a).

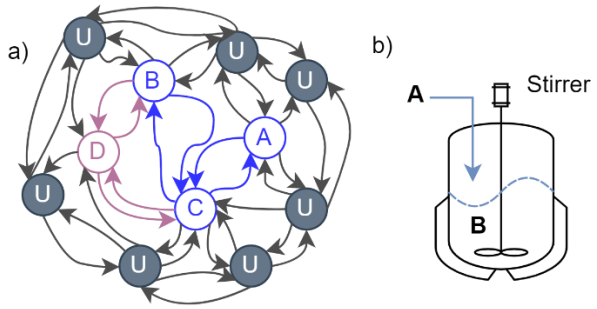

Fig. 2: a): The reactions involved in the studied process. The A, B, C, D are the known chemicals, while the Us are unknown. The blue and purple arrows represent known interactions, and the black arrows are unknown. b): The simplified diagram of SBR mode. B is stored in the vessel, and A is fed at a constant flow rate. The samples were analyzed by Gas Chromatography (GC) to get concentrations of A, B, C in the solution.

Moreover, some chemicals (e.g. Us in Fig. 2a) are not measurable because of the analytics tool limitations. The known reactions and measurable chemicals are modelled via the white-box modelling. The unknown reactions or unmeasurable chemicals are approximated by the black-box modelling. Based on the literature study of similar chemical compounds (Proctor, 2014), the most critical reactions are listed:

R1: A + B ⇌ C

S1: B + C ⇌ D

where A, B, C, D stand for different chemical species, **R1** is the main reaction, **S1** is the side reaction, and ↔ denotes the reversible reaction. The power-law kinetic model is applied and the corresponding reaction rates r [mol/(L·min)] are shown below (Davis and Davis, 2003).

$$r_{R1} = k_{R1,f} C_A C_B - k_{R1,b} C_C \quad \text{Equation 1}$$

$$r_{S1} = k_{S1,f} C_B C_C - k_{S1,b} C_D \quad \text{Equation 2}$$

f and b are forward and backward reactions, respectively. The concentration changes of the chemical components in the system can be represented in the following system:

$$r_A = -r_{R1}, \quad \text{Equation 3}$$

$$r_B = -r_{R1} - r_{S1}, \quad \text{Equation 4}$$

$$r_C = r_{R1} - r_{S1}, \quad \text{Equation 5}$$

$$r_D = r_{S1}. \quad \text{Equation 6}$$

It is noted that the reactions listed in equations above are only known ones in the system. Overall, the reactions in the system are more than two reaction paths. Experimental data are used to approximate other unknown reactions via SI and regression methods. Some chemicals in the mixture cannot be measured or estimated; these are modelled as a lumped group in the SI-M hybrid modelling system. The reaction rate constants k related to temperatures T can be derived from Arrhenius equation:

$$k = A \exp\left(-\frac{Ea}{RT}\right), \quad \text{Equation 7}$$

where A is the pre-exponential factor, Ea is the activation energy [J/mol], and R is the gas constant [J/(K·mol)]. The heat of reversible reactions ΔH [J/mol] can be calculated from the activation energy of the forward reaction Ea_f and the backward reaction Ea_b by:

$$\Delta H = Ea_f - Ea_b. \quad \text{Equation 8}$$

Totally, eight first-principle parameters were required for the two aforementioned reversible reactions, namely pre-exponential factors and activation energies. These reactions and equations are not affected by the type of reactor and they are included in the first-principle model of SI-M (Arrhenius equation block in Fig.1b).

Two types of operations were involved in this paper – SBR and CFR. Fig. 2b gives the simplified diagram of the SBR, and the corresponding experimental setup is Mettler Toledo® Easymax 102 equipped with a dosing system, as shown in Supporting Information. B was firstly stored in the vessel at a constant temperature, and A was then fed at a constant flow rate. The temperatures of the vessels were kept constant. After a specific feeding time (approximately 180 minutes), the solution was kept in the vessel for another 180 minutes for post-reactions. There are two parts for modeling this experiment: feeding and after-feeding (Florit et al., 2018), which were modelled via the following equations:

$$\frac{dC_i}{dt} = r_i - \frac{v_{feed}}{V}(-C_{i,feed} + C_i),\ i = \text{A, B, C, D} \quad \text{Equation 9}$$

$$\frac{dV}{dt} = v_{feed}, \quad \text{Equation 10}$$

$$\frac{dM}{dt} = v_{feed}\rho_{feed}, \quad \text{Equation 11}$$

where C_i is the chemical species concentration [mol/L], v_{feed} is the feed flowrate [L/min], V is the solution volume [L], r is the reaction rate [mol/(L·min)], M is the mass of the solution [g], and ρ_{feed} is the feed density [g/L]. When the feeding stopped, $v_{feed} = 0$ [L/min]. The concentration derivatives are equal to the reaction rates.

RC1 experiment is operated in SBR modes and measures the heat flow rate of reactions. The corresponding experimental setup is provided in Supporting Information. The jacket temperature T_j is measured and dynamically changed, while the temperature inside the reactor T is kept constant. The heat flow rate is calculated based on the energy and mass balances. At first, an adequate amount of B solution in the reactor was heated to the set temperature. Solution A was then added with a constant flow rate. The energy balances over the reactor are

$$Q_{system} = Q_{in} + Q_{reaction} - Q_{out}, \quad \text{Equation 12}$$

$$Q_{system} = MC_p\frac{dT}{dt}, \quad \text{Equation 13}$$

$$Q_{reaction} = V(t)\sum_{m=1}^{2}\Delta H_m r_m(t), \quad \text{Equation 14}$$

$$Q_{in} = Q_{stirrer} + C_{p,dos}(T_{dos} - T)\frac{dM_{dos}}{dt}, \quad \text{Equation 15}$$

$$Q_{out} = UA(T - T_j), \quad \text{Equation 16}$$

where C_p is the specific heat capacity [J/(mol·K)], the subscription dos is for dosage, $Q_{stirrer}$ is the energy from the stirring [W], also considered a baseline, and UA is the product of heat transfer coefficient [W/(m^2·K)] and transfer area [m^2]. The reaction rate r is on unit [mol/(L·s)], and ΔH is the reaction heat [J/mol] (Stoessel, 1997). From the system above, the reaction heat can be estimated.

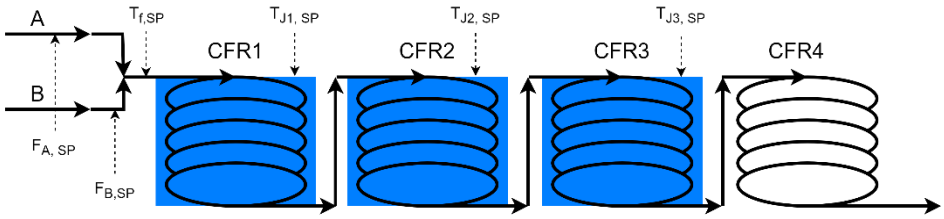

Fig. 3: The simplified diagram of the CFR process in the plant. The streams A and B were forwarded to the first reactor after mixing, and their flowrates are set by $F_{A,SP}$ and $F_{B,SP}$. Only CFR1, 2, 3 have jackets, and the jacket temperatures can be set at the desired values independently with $T_{J1,SP}$, $T_{J2,SP}$, and $T_{J3,SP}$. The last one is cooled via the air at the room temperature.

Fig. 3 shows the CFR process in the plant, consisting of four identical flow reactors with 4.5 L volume each. Two feeding streams containing A and B were first mixed at the pre-set ratios. Then the mixed stream passed through the four reactors one by one. During the whole reaction process, the jacket temperature of each reactor was kept at the desired setpoints. The solution concentrations were analyzed only at the outlet when the plant reached steady states. Each reactor in this process was considered a plug-flow reactor. The steady-state equations for a plug flow reactor without axial and radial dispersion (Smets et al., 2001) can be written as follows:

$$\frac{dC_m}{dV} = \frac{1}{v_{feed}} r_m, \qquad \text{Equation 17}$$

$$\frac{dT}{dV} = \frac{1}{v_{feed}} \left(\frac{\sum_{n=1}^{2} \Delta H_n r_n}{\rho C_p} + \frac{UA(T_j - T)}{\rho C_p V}\right), \qquad \text{Equation 18}$$

where C_m, r_m, and T are component m concentrations [mol/L], component m reaction rates [mol/(L·min)] and temperature [K], V is the reactor volume [L], v_{feed} is the feed flowrate [L/min], ρ and C_p are the density [g/L] and specific heat capacity [J/(mol·K)], ΔH_n and r_n are the reaction heat and rate of reaction n, U, A and T_j are the heat transfer coefficient [W/(m^2·K)], the heat transfer area [m^2], and the jacket temperature [K], respectively. This way, the modelling of the CFRs at a steady state is developed. The four flow reactors in series can represent the process. The heat transfer coefficients of CFR1, CFR2 and CFR3 are already known. An additional heat transfer U of CFR4 is required to determine due to no equipped jacket.

The system modelling SBR, RC1, and CFR constructs the first-principle part. It has the corresponding nine first-priciple parameters (four pre-exponential factors, four activation energies, and one heat transfer coefficient) to be identified during the training.

2.2.2. Hybrid models

The hybrid models were obtained combining dat-driven models with the first principle equations reported in in section 2.2.1. The first-principle equations are with nine parameters: activation energies Ea, pre-exponential factors A and the heat transfer coefficient U. The data-driven part in the hybrid models is based on two polynomial regressors:

$$r_{bb,i}^{1} = (a_i C_A + b_i C_B + c_i C_C + d_i)(-\frac{1}{T} + e_i), \qquad \text{Equation 19}$$

$$r_{bb,i}^{2} = (a_i C_A C_B + b_i C_B C_C + c_i C_C C_A + d_i)(-\frac{1}{T} + e_i), \qquad \text{Equation 20}$$

$$q_{bb} = f\ \sum_{i=1}^{3} r_{bb,i}^{1\ or\ 2}. \qquad \text{Equation 21}$$

where i = A, B, C stands for the three essential chemical components analyzed via GC. $r_{bb,i}^{1}$ is the data-driven part of i reaction rate of the first regression method. C is the concentration in the solution [mol/L]. Furthermore, a, b, c, d, e, f are the corresponding regression coefficients. Overall, the system has sixteen data-driven parameters. The first regression method (see Equation 19) is based on linear regression and additional influence from temperature. In this way, the first-order reactions are assumed. The second regression method (see Equation 20) is from the interaction terms of the second-order polynomial regression.

The hybrid reaction rates [mol/(L·min)] of component i and the hybrid general reaction heat [W] in the system are:

$$r_{hy,i} = r_i + r_{bb,i}^{1\ or\ 2}, \quad \text{Equation 22}$$

$$q_{hy} = \sum_{i=1}^{2} \Delta H_i r_i + q_{bb}. \quad \text{Equation 23}$$

where r_i is the first-principle part from Equation 3, Equation 4, and Equation 5. ΔH is the reaction heat obtained from Equation 8. Together with the ordinary differential equations (Equation 9, Equation 10, Equation 11, Equation 14, Equation 17, Equation 18), the developed hybrid models enable to predict the process outputs (concentration and temperature profiles).

Overall, the hybrid model has twenty-five parameters to be identified (i.e., nine parameters of the first-principle model and sixteen parameters of the statistical model). The objective function values are calculated using the modelled outputs from hybrid models, the experimental outputs, and the parameters. By evaluating the objective values and bounds, SI algorithms can give the parameters for the next iteration. The whole training procedure stops till it reaches the set iterations. As shown in Fig. 1a, this paper considers feed concentrations, feed flow rates, jacket temperatures, and reactor volumes as inputs. The outputs are the concentrations and heat flows. The bounds for first-principle parameters are set according to the literature study, while the data-driven coefficients were searched in the range of [-0.1, 0.1]. The reason why a small range [-0.1, 0.1] was chosen is that the first-principle parts play the primary role in the hybrid modelling to explain the complex reaction processes.

The utilized objective function F in this paper is the sum of Normalized Mean Squared Errors (NMSEs) with L1 regularization term, defined as:

$$\begin{aligned} F &= \sum(NMSEs) + L1, \\ &= \sum_{k}^{l} \frac{\sum_{i=1}^{n^k} \left(\frac{w(\hat{y}_i^k - y_i^k)}{\max(y^k) - \min(y^k)} \right)^2}{n^k} + \lambda \sum_{j=1}^{m} |\theta_j| \end{aligned} \quad \text{Equation 24}$$

where n is the number of the training data points, $\hat{y}^k$ is the predicted output of variable k, y is the experimental data, and w is the adjustment factor, set to 10 in this paper. For the L1 regularization term, $\lambda = 10^{-3}$ and θ_j is the j-th parameter among the m hybrid model parameters. SI algorithms searched within the boundaries via algorithms inspired by mold, birds, and ants. The optimal set of parameters is given by minimizing the objective value F.

This paper studies two regression methods (i.e., Equation 19 and Equation 20) and three SI algorithms (i.e., SMA, ACO, PSO in section 2.1) in SI-M. The simulation and training routines are implemented in Python. They are also compared with a commonly used gradient descent method, L-BFGS-B (Limited-memory Broyden–Fletcher–Goldfarb–Shanno with Bounds) implemented in SciPy (Byrd et al., 1995; Machalek et al., 2021; Morales and Nocedal, 2011). The calculation used a machine having Intel®Xeon®Gold 5220 CPU @ 2.20GHz, 2195 Mhz, 18 Core(s), 36 Logical Processor(s).

2.3. Process optimization

Using the trained hybrid models, SI-O (Swarm Intelligence-based Optimization) was conducted to maximize the productivity of C. The developed optimization agent provides the optimal setpoints of A or B inlet flow rate ($F_{A,SP}, F_{B,SP}$), feed temperature ($T_{f,SP}$), and the jacket temperatures of the first three reactors ($T_{J1,SP}, T_{J2,SP}, T_{J3,SP}$) as shown in Fig. 4.

The optimization agent relies on the developed model and the algorithms to update the setpoints. The initial candidates, namely starting setpoints in this study, can be randomly generated or assigned by users. Based on the given candidates, the process can be simulated. The corresponding modelling results are evaluated according to the C productivity. In this process, the price of product C is much higher than the raw materials and utilities. The constraints to avoid the plant shut-down are also included in the evaluation function, where the evaluation value would be -1 if the constraints cannot be maintained. The working flow of the optimization stops when the defined requirements are met, namely the number of iterations. If the requirements are not met, the candidate setpoints are updated by SI algorithms, and it repeats the modelling and evaluation steps.

The optimization strategy of the applied process is slightly different from normal operations in order to decrease the calculation times of the agent. The calculation time is critical in this continuous process, and the product is lost when the process is waiting for the new optimal setpoints. Therefore, a modified optimization strategy in SI-O is proposed and illustrated in Fig. 5. It recclculates the process setpoints (a) from the agent unless the process inputs (u) are different from the previous optimal history. If the deviation between the new $u(t+1)$ and the previous optimal $u_{opt}(t_{opt})$ is low, the reward function r is calculated based on the current states $s(t)$, historical actions $a(t_{opt})$ and the corresponding inputs $u(t+1)$. If the reward is higher than 0, the new setpoints $a(t+1)$ are $a(t_{opt})$. Otherwise, the optimization agent is applied to get the new setpoints a_{agent}. The system stops when the time t_{end} is reached.

As methined above, the optimization strategy highly depends on the developed process models. If the model cannot represent the process anymore, the proposed setpoints would be unsuitable. Fortunately, the process is already proven to operate safely at some fixed setpoints. With these setpoints, the maximum productivity cannot be achieved, but the process can avoid unexpected shut-downs. In order to ensure the process keep working, a back-up mode can be added. The optimization mode is applied initially from the beginning of the process. The corresponding plant outputs are compared with the model outputs to check if the model is still valid or not. If the model is invalid, the backup mode with fixed setpoints is employed. The plant outputs from the backup mode are also compared to the previous plant outputs from the optimization mode. Suppose the performance of the optimization mode is better than the backup mode. In that case, the optimization mode will be reused, meaning the process behaviours still follow the model behaviours even though the model cannot provide accurate absolute values. Otherwise, the backup mode will be used at the following time of the plant operation, and the model is required to be retrained. A graphical user interface for SI-M/O is designed to have a clear view of its performance, which is implemented in python via Tkinter (Lundh, 1999). It consists of four sections: process information, settings, running messages and results.

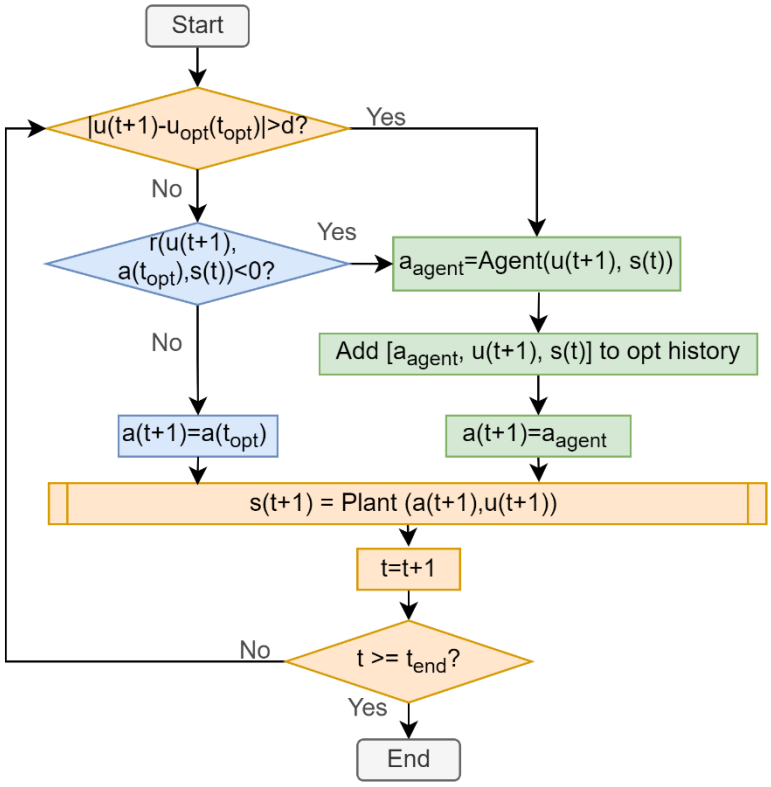

Fig. 5: The SI-O optimization strategy of the process, where u(t) is the inputs of the plant at time t, r is the defined reward function, a is the setpoints applied to the process, s is the states of the process, and d is the set deviation.

3. Results

3.1. Process modelling

With the methodology SI-O explained in section 2.2, the hybrid model for complex reaction processes can be developed. A total of sixteen approaches are investigated in this paper, resulting from the involved two sub-training categories (one-time and step-by-step), two regression methods (Regression 1 in Equation 19 and Regression 2 in Equation 20), and four optimization algorithms (ACO, PSO, SMA, and L-BFGS-B). The following section gives the performance analysis of different approaches. The sensitivity analysis and the applications to the production plant are also given based on the optimal approach.

3.1.1. Performance analysis

All sixteen approaches are trained ten times with a maximum of 600 iterations. The using boundaries are the same, which are [0, 1] for the first-principle parameters including pre-exponential factors A, activation energies Ea, and heat transfer coefficients U with scaling factors [1.5e4, 1e0, 1e4, 1e0, -1e4, -1e4, -1e4, -1e5, 1e1], and the range [-0.1, 0.1] was employed for the data-driven parameters. The number of solution candidates is 30 for SI algorithms (ACO, PSO, SMA). One of the advantages of SI is that no initial guessed point is required. However, the initial starting points of the L-BFGS-B method are compulsory, set to the random values between the upper and lower bounds in each run.

In this paper, L-BFGS-B returns the highest objective values with the magnitude of 10^5 in every training approach. That is because L-BFGS-B can easily converge to the local optimum; most probably this is related to the random initial conditions adopted in this paper. Also, the used functions are with high nonlinearity. The ideal objective functions of this technique are usually required to be continuous and easily differentiable, which is not the case in this paper. A high amount of extra fine-tuning work is required in L-BFGS-B for a fair comparison. Hence, it is suggested to apply L-BFGS-B when the starting solution has high confidence and the used functions are not highly nonlinear.

The average objective values F_{avg} at the end of the training and the average calculation time t_{cal} in seconds are presented in Table 1.

Table 1: The average objective value F_{avg} and average calculation time t_{cal} in seconds for ten runs. The number in [ ] are the approach index as shown in Fig. 6.

SI algorithms	Regression 1				Regression 2			
	One-time		Step-by-step		One-time		Step-by-step	
	F_{avg}	t_{cal}	F_{avg}	t_{cal}	F_{avg}	t_{cal}	F_{avg}	t_{cal}
ACO	46.5 [1]	1,730.4 [1]	41.9 [2]	1,784.9 [2]	54.9 [3]	2,057.0 [3]	39.3 [4]	1,304.1 [4]
PSO	53.4 [5]	20,130.4 [5]	46.7 [6]	11,512.9 [6]	54.8 [7]	14,122.5 [7]	52.1 [8]	11,227.2 [8]
SMA	45.0 [9]	20,524.4 [9]	**36.2 [10]**	**7,053.4 [10]**	55.4 [11]	20,038.0 [11]	40.3 [12]	7,733.0 [12]

The average objective values F_{avg} for SI approaches are all below 60. The approaches with ACO take less calculation time t_{cal} than the others, and their average objective values F_{avg} are lower than those with PSO. The PSO approaches achieve higher F_{avg} values than ACO and SMA, and their calculation time is higher than those in the same sub-categories except SMA + one-time approaches. For SMA, the one-time approaches take more calculation time than ACO and PSO, and the step-by-step approaches consume less than PSO but more than ACO. Also, the lowest average objective value F_{avg} is achieved by Approach 10: SMA + Regression 1 + step-by-step.

Generally, the step-by-step approaches achieved lower F_{avg} and lower calculation time than the one-time approaches. The time differences between step-by-step and one-time for SMA are high as close to three times. The reason is that the step-by-step approaches splits a high-dimension task with twenty-five parameters into two sub-tasks: one is for nine first-pricinple parameters and the other is for sixteen data-driven parameters. In conclusion, Approach 10: SMA + Regression 1 + step-by-step achieved the lowest objective value with a reasonable calculation time of 7,053.4 seconds.

For the performance analysis of different modelling approaches in ten training runs, Fig. 6 gives the boxplot where the boxes and lines depict the distributions of final objective values at the end of training. The soild box indicates the range of the central 50% of the data, with a central white line marking the median values. The bounds of boxes are defined when greater or less than 25% of the data. The extended lines from each box capture the remaining data range, with circles indicating outliers. It can be seen that Approach 4, Approach 10 and Approach 12 have most objective values lower than 40. Among the three approaches, Approach 10 outperforms the others showing no outliers, a low variance and the lowest median.

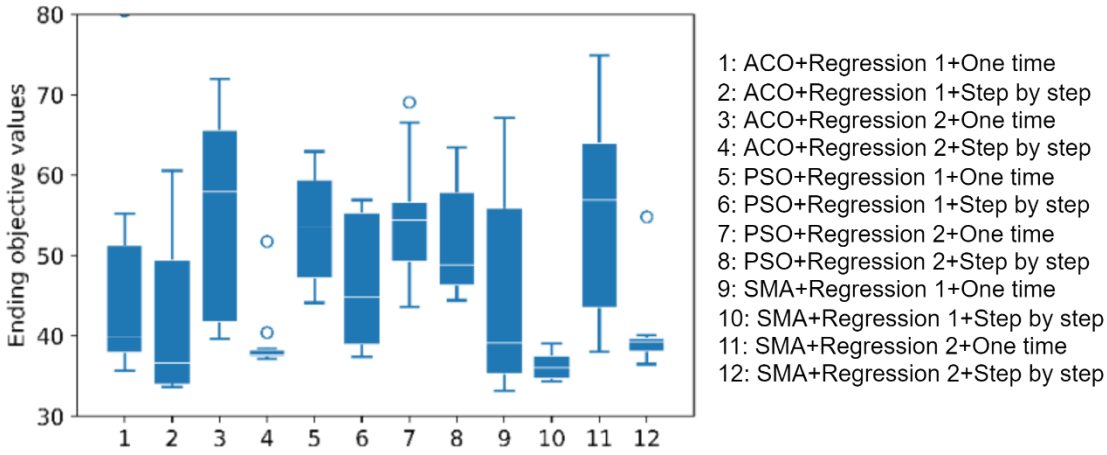

Fig. 6: Box plot of the final objective values via SI algorithms. The boxes and lines depict the distributions of final objective values at the end of ten runs for various modelling approaches. The box indicates the range of the central 50% of the data, with a white line marking the median values. The extended lines from each box capture the range of the remaining data, with circles indicating outliers.

3.1.2. Searching progress of swarm intelligence algorithms

This section discusses how the various SI algorithms perform the parameters searching while training the model. The training processes of the twelf approaches are shown in Fig. 7. In this Figure, a, b, and c plots are for ACO, PSO, and SMA, respectively. The dashed lines are the objective value minimum during the 600 iterations, and the dotted lines are the maximum.

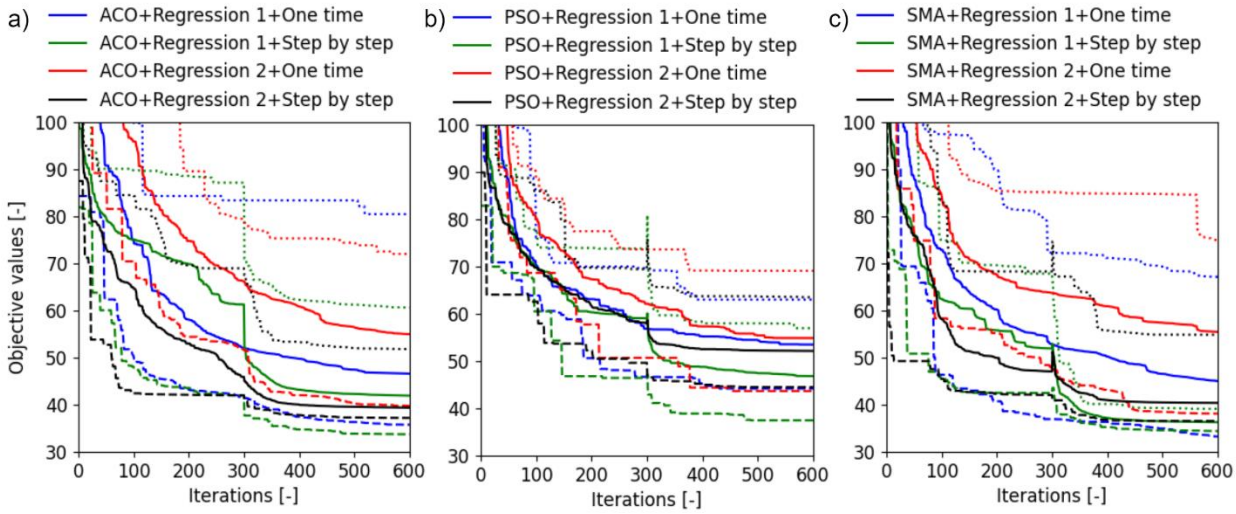

Fig. 7: The training processes of 12 approaches with swarm intelligence algorithms: a) ACO, b) PSO, c) SMA. The solid lines are the average objective values. The dashed lines are the objective value minimum during the 600 iterations, and the dotted lines are the objective value maximum during the 600 iterations.

For all ACO, PSO, and SMA approaches, step-by-step ones (black and green lines in Fig. 7) achieve lower average objective values than one-time approaches (blue and red lines in Fig. 7). In the step-by-step apporoch, the first 300 iterations are dedicated for the first-principle models. Then, the first-principle parameters are fixed and the other 300 iterations are for data-driven parameters. The improvements are significant when the training starts over at the 300th iteration. For some runs after the 300th iteration in SMA

and PSO, the objective values increase slightly because the primary search does not find the desired direction. The difference between ACO and SMA approaches is minor, but the PSO approaches are usually with higher ending objective values than the other two. Even the maximum objective values of SMA + Regression 1 + step-by-step are smaller than those of most PSO approaches. In sum, Approach 10: SMA+Regression 1 + step-by-step gets the lowest mean objective value. Also, its variance, which can be indicated from the difference between the minimum and the maximum runs, is smaller than the others.

The reasons why different SI algorithms perform differently are investigated too. Three randomly selected candidates among thirty are plotted in Fig. 8, indicating their position changes during the training process. The randomly chosen parameters are Parameter 9 (one first-principle parameter) and Parameter 10 (one data-driven parameter). The arrows in Fig. 8 show the changing directions of the candidates in 600 iterations. The training process of other example parameter sets can be found in Supporting Information.

The ACO approaches show fewer movements than PSO and SMA, resulting in low stability, as seen in Fig. 6. After several movements, the candidates arrive at similar positions. The first-principle parameters (e.g., Parameter 9) keep the same after the 300^{th} iteration in step-by-step approaches, resulting in vertical arrows. PSO-based approaches, which cost high computational time, explore the entire region within the boundaries. However, the final solutions of PSO approaches within the 600 iterations are with high objective values. The reasons can be that the updates of PSO candidates do not take much into account the information from the other candidates, resulting in the candidates working more independently than the other approaches. The SMA-based approaches balance the exploration and the computational cost appropriately. The behaviours of condidates in Fig. 8 explain why Approach 10: SMA + Regression 1 + step-by-step outperforms the other approaches.

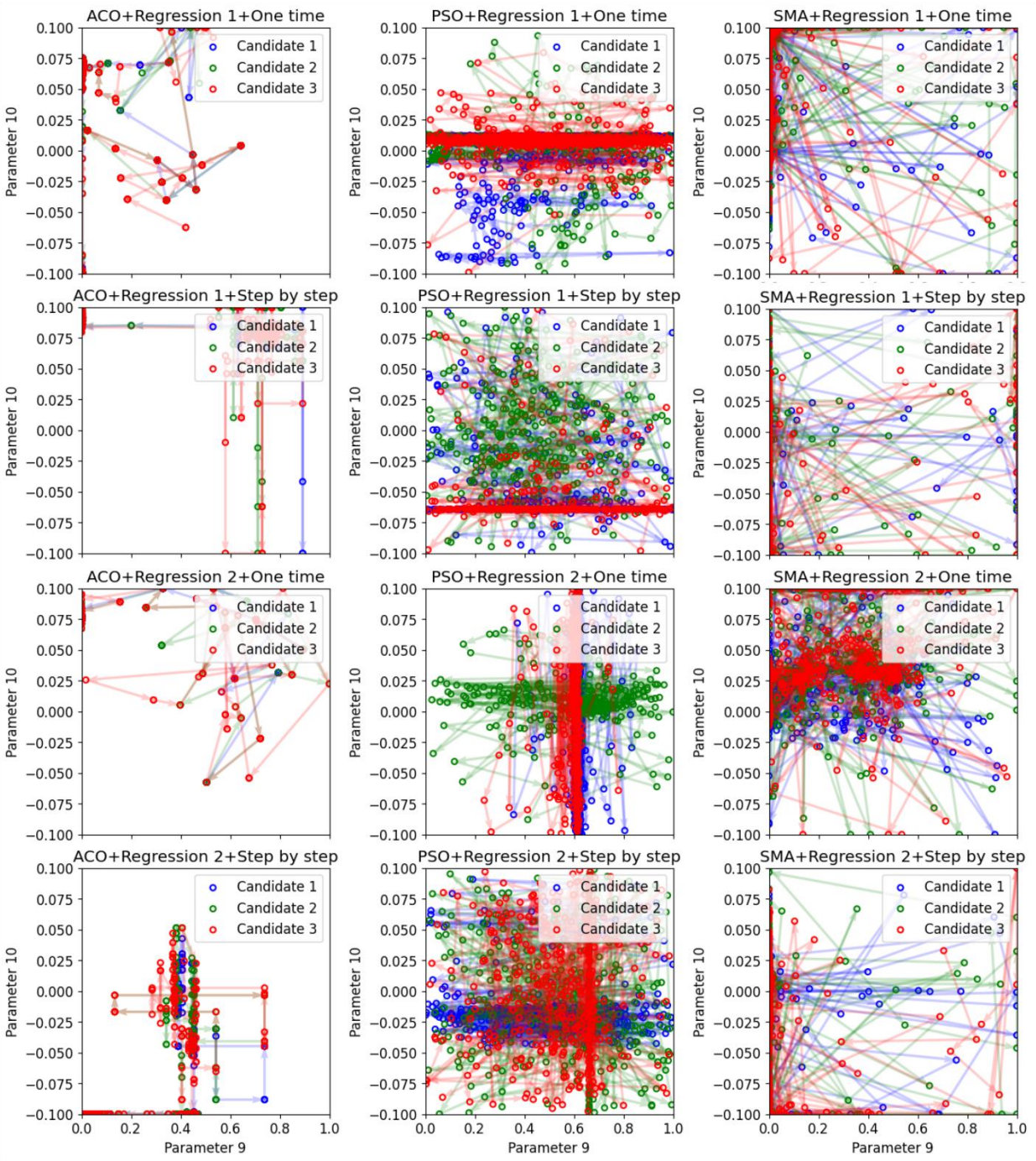

Fig. 8: The position changes during the training process for Parameter 9 (a white-box parameter) and Parameter 10 (a black-box parameter). Three candidates (blue, green, red) are involved among 30 candidates, and the corresponding arrows show the changing directions of the candidates in 600 iterations.

3.1.3. Model validation and sensitivity analysis

The model developed via Approach 10: SMA + Regression 1 + step-by-step is used to simulate the steady states of the plant process (see Fig. 3). Fig. 9 gives the corresponding model performance through two tests.

The scatters are experimental data points, and the lines are modelling results. The experimental data is limited because the composition analysis is only at the exit of the fourth continuous flow reactor (CFR) and the temperature measurements are only available at the exit of the four CFRs.

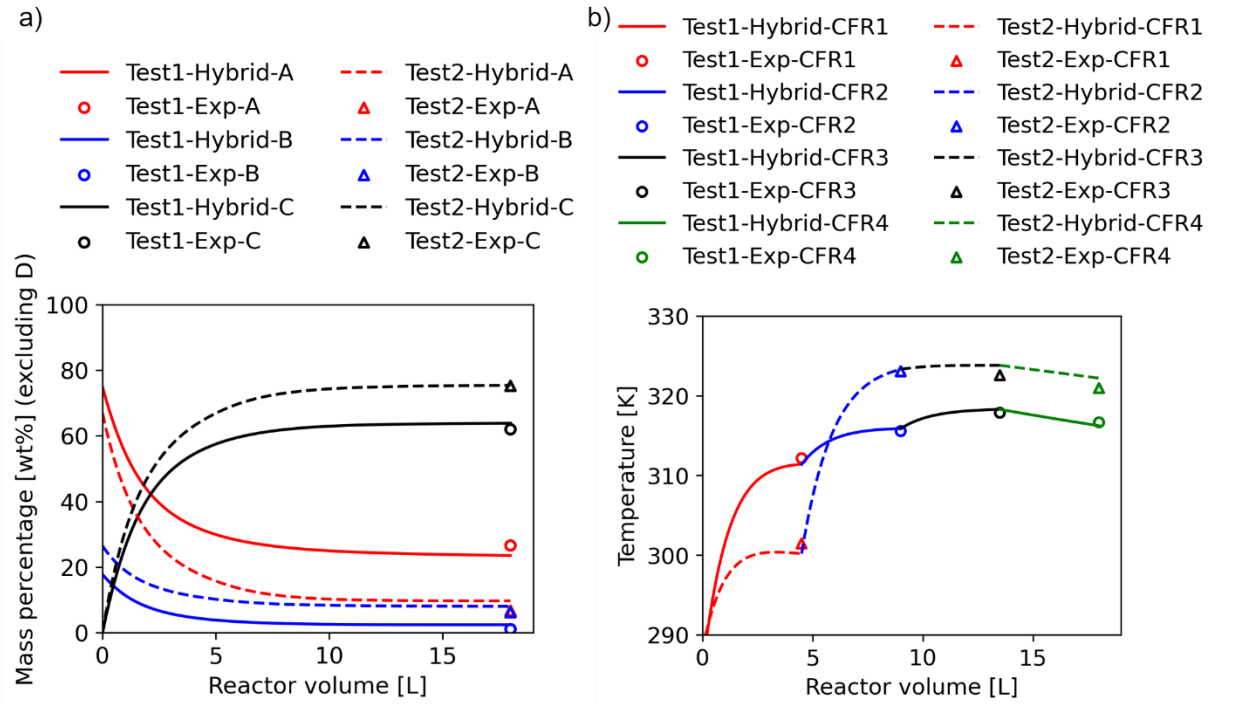

Fig. 9: Model validation of SMA + Regression 1 + step-by-step method via two tests(Test1:solid lines, Test2: dashed lines). a): Normalized concentrations in weight mass percentage [wt%] of components A (red), B(blue) and C(black). The scatters are the experimental data at the exit of the fourth reactor. The lines are the modelling results from the developed model; b): Temperatures of continuous flow reactor1 (CFR1)(red), CFR2(blue), CFR3(black), and CFR4(green). The scatters are the experimental data at the exit of four reactors. The lines are the modelling results from the developed model.

These two tests used different plant settings compared to the training data set, which are listed in Table 2, while the concentrations of feeding raw materials A and B are the same, namely 40.5% and 91.9%, respectively. F_A and F_B are the feed flow rates of A and B in [kg/h], T_f is the feeding temperature in [°C], and T_{J1}, T_{J2} and T_{J3} are the jacket temperatures in [°C].

Table 2:The six setpoints of the plant for validating the developed model. F_A and F_B are the feed flowrates of A and B, T_f is the feeding temperature, and T_{J1}, T_{J2} and T_{J3} are the jacket temperatures.

	$F_{A,SP}$ [kg/h]	$F_{B,SP}$ [kg/h]	$T_{f,SP}$ [°C]	$T_{J1,SP}$ [°C]	$T_{J2,SP}$ [°C]	$T_{J2,SP}$ [°C]
Test1	101.2	54.5	12.5	37.5	42.5	45.0
Test2	82.1	73.6	15.0	25.0	50.0	50.0

It can be seen in Fig. 9 that the model predicts the plant process accurately in two different tests. In Test1, the starting B mass percentage is less than in Test2, leading to the lower ending B, C and the higher ending A. The jacket temperatures highly influence the temperatures inside the reactors. The higher jacket temperatures result in higher exit temperatures. The exit temperature of the fourth reactor is slightly lower than the previous reactor because the fourth CFR does not have a jacket. CFR4 is influenced by the room

temperature, which keeps at 20 °C in the plant. The developed model with Approach 10: SMA + Regression 1 + step-by-step successfully simulates the plant process at steady states with high accuracy.

As stated in section 2.2, the involved parameters play important roles in the developed model. Therefore, a sensitivity analysis was conducted to check the influence of parameters on the obtained model. The used reference settings are from Test1, including the A and B feeding flow rates, the feeding temperature, and the jacket temperatures. Based on the obtained parameters, the changes in parameters are in the ranges of 10%, 20% and 30%, respectively. 1,000 parameter samples are randomly generated using Latin Hypercube Sampling (LHS). The corresponding concentration and temperature profiles are shown in Fig. 10, where the filled areas illustrate the model vibrations resulting from the changing parameters. The solid line is the reference model performance for Test1 with the obtained parameters via Approach 10: SMA + Regression 1 + step-by-step.

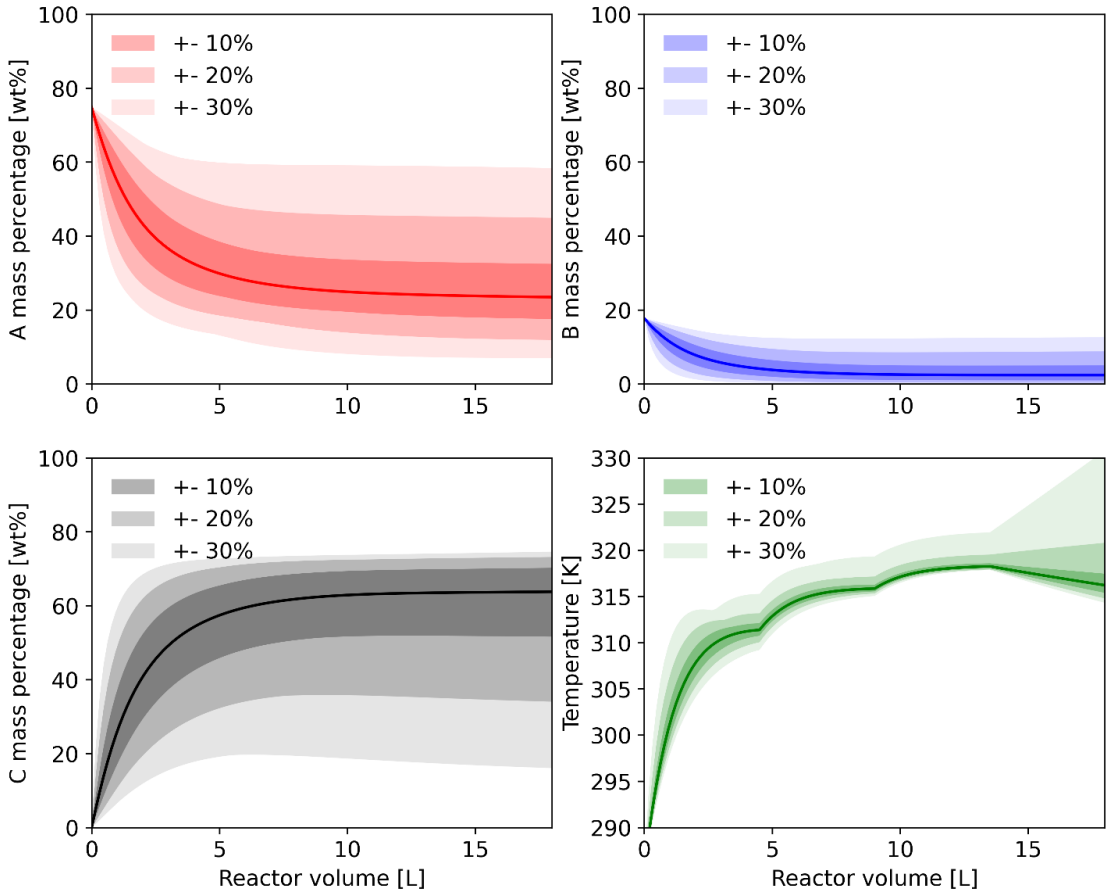

Fig. 10: The changes of parameter values in 10%, 20% and 30% to the model in Test1. LHS within selected boundaries randomly generates the changes in the parameters in 1,000 samples. The solid line is based on the model with the obtained optimal parameters.

The changes in the parameters in the range of 10% significantly influence the concentration profiles but have much less influence on the temperature profiles, resulting from that the major number of parameters in the model is related to reaction kinetics. The parameters connected to the reactors are mainly from the available database, instead of estimated via the SI-M training. The parameter changes within 30% highly affect the model performance both for concentration and temperature profiles. The changes higher than 30% lead to the failure of the model convergence. In conclusion, the developed model is sensitive to the involved

parameters. It can explain why the conventional algorithm (e.g. L-BFGS-B in this paper) obtained high objective values when training the model, because the inappropriate starting parameters can also fail the model convergence. However, the SI algorithms including ACO, SMA and PSO can easily overcome this shortcoming, benefiting from many beginning searching candidates. Compared to conventional algorithms, the risk of inappropriate starting candidates or parameter ranges is reduced via SI-M.

3.2. Process optimization

Based on Approach 10: SMA + Regression 1 + step-by-step, the model of the plant process (see Fig. 3) is developed and the parameter sensitivity analysis is performed. The following process optimization is based on the obtained model, and the six setpoints are changed according to the raw material changes via SI-O (Swarm Intelligence-based Optimisation). This section presents the performance analysis of SI-O. Furthermore, the results of experimental plant test via SI-O is given. In addition, a graphical user interface is designed.

3.2.1. Performance analysis

The aim of process optimization is to maximize the productivity and to ensure the plant continuous operation. The corresponding reward function is designed based on these requirements. The constraints can be divided into Hard Constraints (HCs) and Soft Constraints (SCs). In this study, there are two hard constraints to avoid the plant shut-down and to operate the plant safely, namely 1) the feeding flowrate of raw materials A and B should be smaller than the maximum values set by the operators; 2) the total flow rate should be smaller than the maximum value given by the operators as well. Moreover, two soft constraints should be obeyed based on the human expert experience for reactions and reactors: 1) the residence time should be in the range of 4 and 7 minutes; 2) the feeding ratio of pure material B vs A should be lower than 1.25 and higher than 0.75. Therefore, the final reward function R is

$$R = \begin{cases} -1, & \text{if HCs are not met} \\ wt_c M - w_{rt} P_{rt} - w_{ratio} P_{ratio}, & \text{if HCs are met} \end{cases} \quad \text{Equation 25}$$

where wt_c is the weight mass percentage (%) of the product, w_{rt} and w_{ratio} are the weights assigned to the punishments of the residence time (P_{rt}) and the feeding ratio (P_{ratio}), respectively. The punishments P_{rt} and P_{ratio} are calculated by the absolute difference from the desired ranges, and the corresponding w_{rt} equals 1.0 and w_{ratio} is 10.0.

ACO, PSO, SMA and L-BFGS-B are employed to obtain the maximum reward. Their performances according to the changes of the feeding raw material A and B weight percentage concentrations are shown in Fig. 11. The x-axis represents 10 different changing steps of raw materials A and B. Fig. 14a gives how the feed concentrations of raw materials A and B change. It can be seen that the changes are not gradually but in steps. In this study, the step values can keep for a few hours, which is much longer than the residence time in the magnitude of minutes. That is why the steady-state model of the plant process can be utilized in SI-O. Fig. 14b and Fig. 14c illustrate the averages and standard deviations of reward values for 10 runs according to raw material changes in Fig. 14a, where the circles are the average values, and the vertical lines are the standard deviations.

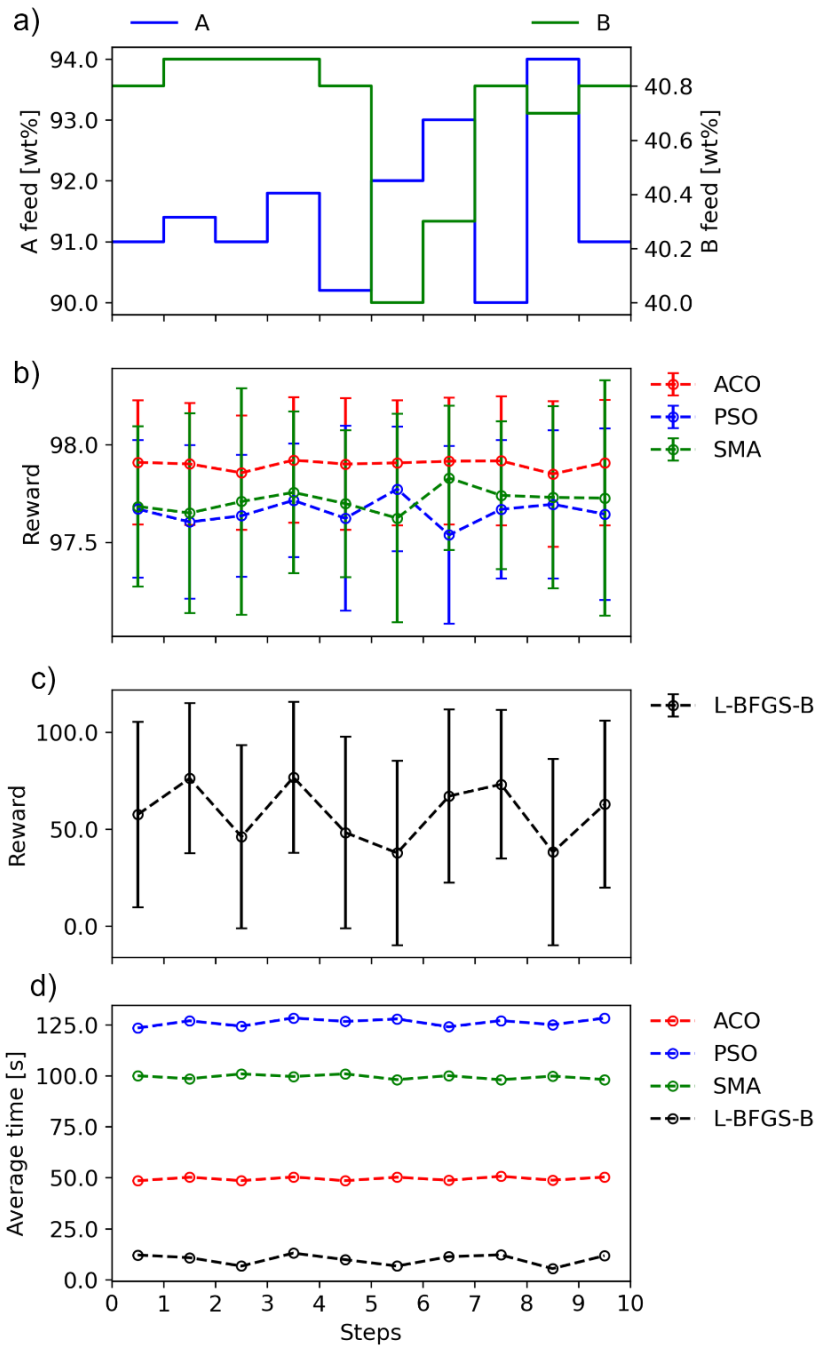

Fig. 11: The performance of optimization agents with different swarm intelligence algorithms and L-BFGS-B, the x-axis is standing for 10 different changing steps of raw material A and B. a): The changes of the feed concentrations of raw material A and B; b) and c): The averages and standard deviations of reward values for 10 runs according to raw material changes(the circles are the averages, the vertical lines are the standard deviations); d): The average calculation times for 10 runs according to raw material changes.

ACO got the highest reward average values and the lowest standard deviations of ten runs for the ten different changing steps. In contrast, the average reward values and the standard deviations of PSO and SMA are close. It is noted that in SI-M, Approach 10 with SMA obtained the best performance while ACO achieved the best performance in SI-O. One of the reasons is that ACO is more suitable to low-dimension

tasks, such as six setpoints in SI-O compared to twenty-five parameters in SI-M. It can get the optima much faster than SMA and PSO, which is crucial because the delay in applying new setpoints causes the ongoing product losses in the process. However, ACO shows high variance and low stability when it comes with high-dimension problems with over twenty parameters. The employed functions, which are the objective function F (see Equation 24) in SI-M and the reward function R (see Equation 25) in SI-O have different characteristics. The objective function F is continuous and considering the regularization term, but the reward function R is discontinuous and no regularization term is involoved.

The performance of L-BFGS-B in SI-O is similar to its performance in SI-M. It easily traps in the local optimal point. It sometimes even cannot find a positive reward value in SI-O, leading to the lowest average value and the highest standard deviations. However, it can be seen from Fig. 14d that the average calculation times for L-BFGS-B is the least. L-BFGS-B can be an applicable algorithm if the starting point and the applied boundaries are with high condfidence. It has to be emphasized that in this research, the starting point is randomized, and the applied boundaries are the same as the other three SI algorithms. It also shows one of the advantages of SI algorithms is that less work is required to tune the starting points and searching boundaries. ACO used the least calculation average time and the highest average reward values among the three SI algorithms. Hence, ACO is employed for the further study of SI-O. An experimental plant test with SI-O was conducted, and the results are provided in the following section.

3.2.2. Experimental and simulation tests

In the experimental tests for SI-O, the conditions of the plant process were kept the same, namely, the concentrations of feeding raw materials A and B were constant, and the maximum flow rates set by operators were constant as well. The analyzed samples of the outlet stream were taken when the process reached steady states. The "Current approach" used the setpoints given by the company internal calculations, including the feeding flowrates of A and B ($F_{A,SP}, F_{B,SP}$), the feed temperatures ($T_{f,SP}$), and the jacket temperature of the first three reactors ($T_{J1,SP}, T_{J2,SP}, T_{J3,SP}$) as illustrated in Fig. 3. The "Opt. temperatures" applied the optimal temperatures ($T_{f,SP}, T_{J1,SP}, T_{J2,SP}, T_{J3,SP}$) only from the optimization agent with ACO but used the same flow rates as the current approach. The reason for applying only the optimal temperatures is that, based on the expert experience, the temperature influence might be less than the impact of the flow rates. This way, the plant will not shut down unexpectedly at the first stage. The second stage, "Opt. temperatures & flowrates", applied both the optimal temperatures and the optimal flowrates ($T_{f,SP}, T_{J1,SP}, T_{J2,SP}, T_{J3,SP}, F_{A,SP}, F_{B,SP}$). The experimental test results of the normalized C productivity for three approaches are given in Fig. 12.

The colourful bars are the average normalized productivities in [kg/h], and the solid lines are the standard deviations of taken samples. It can be seen that with only the optimal temperatures ($T_{f,SP}, T_{J1,SP}, T_{J2,SP}, T_{J3,SP}$), the productivity increased by 3.5%. The standard deviations of analyzed samples in "Opt. temperatures" are higher than the other two because the limited three samples were taken and analyzed. With the optimal temperatures and flowrates ($T_{f,SP}, T_{J1,SP}, T_{J2,SP}, T_{J3,SP}, F_{A,SP}, F_{B,SP}$), the mean productivity can increase by 5.3% compared to the current approach. It can be seen that the experimental test with SI-O in the plant was successful.

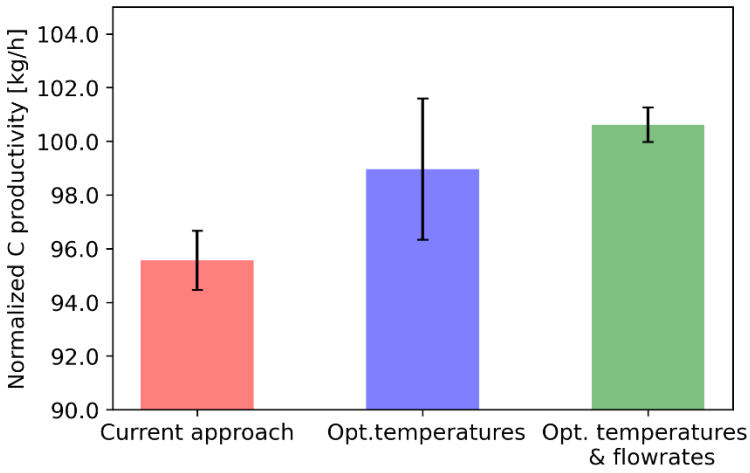

Fig. 12: The plant test results based on the optimization agent with ACO. The bars represent the average normalized C productivity in kg/h, and the solid lines are the standard deviations of taken samples. The "Current approach" used the setpoints given by the internal calculations from the company. "Opt. temperatures" applied only the optimal temperatures from the optimization agent with ACO. "Opt. temperatures & flowrates" utilised both the optimal temperatures and the optimal flow rates from the agent.

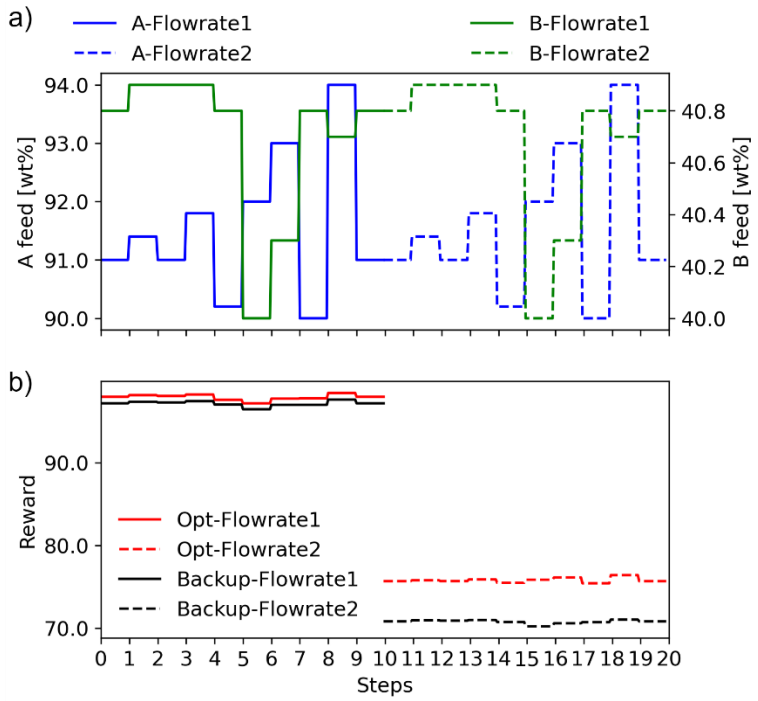

Fig. 13: The performance of SI-O and backup mode for different hard constraints, where "Flowrate1" (solid lines) is 132 kg/h for the maximum flowrate of A, and the "Flowrate2" (dashed lines) is 70 kg/h. a): The changes in the feed concentrations of raw materials A and B; b): The obtained reward values of optimization and backup modes according to raw material changes.

In the previous test, the raw material concentrations are constant, and the maximum flow rates of A and B and the maximum total flow rates are also constant. The maximum flowrates are the part of hard constraints, as shown in Equation 26. To check the performance of SI-O for different hard constraints, the tests in the simulation environment are also conducted, and the results are shown in Fig. 13. Normally, the maximum flowrate should be 132 kg/h, as "Flowrate1" (solid lines). However, this limitation can be 70 kg/h due to the pump failures or the requirements of the downstream processes, as "Flowrate2" (dashed lines). It can be seen in Fig. 13 that SI-O always gives a higher reward value than the backup mode which used fixed setpoints. However, the improvement for "Flowrate1" is much less than "Flowrate2" because the fixed setpoints has been optimized for "Flowrate1" in the previous years. However, the backup mode cannot handle the new "Flowrate2", where SI-O shows the significant improvement. The proposed optimization strategy SI-O obtains the high flexibility, whether the maximum flow rate of A is low or high. At the end, the graphical user interface for SI-M/O was designed and implemented in python, as shown in Supporting Information.

4. Conclusions

This study proposes "SI-M/O" a Swarm Intelligence-based Modelling and Optimization structure to model and optimizate complex reaction processes. The proposed methodology gives a more promising result than the commonly used L-BFGS-B (Limited-memory Broyden–Fletcher–Goldfarb–Shanno with Bounds). Due to the used functions and unavailable starting points, the approaches with L-BFGS-B are unsuitable in this case. In SI-O, Slime Mold Algorithm (SMA) outperforms Ant Colony Optimization (ACO), and Particle Swarm Optimization (PSO) integrated with the residual regression term with the first-order regression (Regression 1) and the step-by-step training. This study fills in the empty part of SMA applications in chemical engineering. With the help of SI-M (Swarm Intelligence-based Modelling), the obstacles of unknown reactions and limited data in modelling complex reaction processes are overcome, obtaining the models with high accuracy and reasonable calculation time. The SMA + Regression 1 + step-by-step approach shows the advantages with reasonable calculation time and lower objective values than others. This hybrid modelling structure can also be applied to other processes with similar problems, such as for unexpected impurities.

According to the SI-M developed models, the process optimization for the plant was developed based on the proposed SI-O (Swarm Intelligence-based Optimization). The SI-O strategy with ACO gives the highest average reward values with the second lowest calculation time. It also increased productivity by 5.8% compared to the current approach in the experimental production plant test. The optimization mode in SI-O always gives higher reward values than the backup mode in the simulation. Moreover, the plant can operate continuously through the accompanied backup mode even when the models are not valid longer. This approach can also be applied other continuous processes, e.g. continuous distillation and continuous crystallization.

Future work will apply the proposed SI-M to obtain a hybrid dynamic model of a production plant. The appliablilty of SI-M can also be evaluated. The current SI-O is based on the steady-state models. The future work will apply the dynamic models in SI-O as well to investigate its performance compared to advanced machine learning techniques, such as reinfrocement learning.

Acknowledgements

The authors acknowledge the funding from VLAIO, DAP^{2}CHEM: Real-time data-assisted process development and production in chemical applications (HBC.2020.2455).

References

Abdel-Basset, M., Mohamed, R., Chakrabortty, R.K., Ryan, M.J., Mirjalili, S., 2021. An efficient binary slime mould algorithm integrated with a novel attacking-feeding strategy for feature selection. Comput. Ind. Eng. 153, 107078. https://doi.org/10.1016/j.cie.2020.107078

Ammar, Y., Cognet, P., Cabassud, M., 2021. ANN for hybrid modelling of batch and fed-batch chemical reactors. Chem. Eng. Sci. 237, 116522. https://doi.org/10.1016/j.ces.2021.116522

Anagnostopoulos, I., Zeadally, S., Exposito, E., 2016. Handling big data: research challenges and future directions. J. Supercomput. 72, 1494–1516. https://doi.org/10.1007/s11227-016-1677-z

Babanezhad, M., Behroyan, I., Nakhjiri, A.T., Marjani, A., Rezakazemi, M., Shirazian, S., 2020. High-performance hybrid modeling chemical reactors using differential evolution based fuzzy inference system. Sci. Rep. 10, 1–11. https://doi.org/10.1038/s41598-020-78277-3

Blum, C., 2005. Ant colony optimization: Introduction and recent trends. Phys. Life Rev. 2, 353–373. https://doi.org/10.1016/j.plrev.2005.10.001

Byrd, R., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. J. Sci. Comput. 16, 1190–1208.

Chakraborty, A., Kar, A.K., 2017. Swarm Intelligence: A Review of Algorithms, in: Nature-Inspired Computing and Optimization. pp. 475–494. https://doi.org/10.1007/978-3-319-50920-4

Davis, M.E., Davis, R.J., 2003. The Basics of Reaction Kinetics for Chemical Reaction Engineering, in: Fundamentals of Chemical Reaction Engineering. pp. 1–52.

Fang, H., Zhou, J., Wang, Z., Qiu, Z., Sun, Y., Lin, Y., Chen, K., Zhou, X., 2021. Hybrid method integrating machine learning and particle swarm optimization for smart chemical process operations. Front. Chem. Sci. Eng.

Feyo De Azevedo, S., Dahm, B., Oliveira, F.R., 1997. Hybrid modelling of biochemical processes: A comparison with the conventional approach. Comput. Chem. Eng. 21. https://doi.org/10.1016/s0098-1354(97)87593-x

Florit, F., Busini, V., Storti, G., Rota, R., 2018. From semi-batch to continuous tubular reactors: A kinetics-free approach. Chem. Eng. J. 354, 1007–1017. https://doi.org/10.1016/j.cej.2018.08.044

Hassanpour, H., Mhaskar, P., House, J.M., Salsbury, T.I., 2020. A hybrid modeling approach integrating first-principles knowledge with statistical methods for fault detection in HVAC systems. Comput. Chem. Eng. 142, 107022. https://doi.org/10.1016/j.compchemeng.2020.107022

Kennedy, J., 2006. Swarm Intelligence, in: Handbook of Nature-Inspired and Innovative Computing Integrating Classical Models with Emerging Technologies. pp. 187–220.

Kennedy, J., Eberhart, R., 1995. Particle Swarm Optimisation. Proc. ICNN'95 - Int. Conf. Neural Networks 4, 1942–1948. https://doi.org/10.1109/ICNN.1995.488968

Kumar, C., Raj, T. Dharma, Premkumar, M., Raj, T. Dhanesh, 2020. A new stochastic slime mould optimization algorithm for the estimation of solar photovoltaic cell parameters. Optik (Stuttg). 223, 165277. https://doi.org/10.1016/j.ijleo.2020.165277

Li, S., Chen, H., Wang, M., Asghar, A., Mirjalili, S., 2020. Slime mould algorithm : A new method for stochastic optimization. Futur. Gener. Comput. Syst. 111, 300–323. https://doi.org/10.1016/j.future.2020.03.055

Lundh, F., 1999. An introduction to tkinter.

Machalek, D., Quah, T., Powell, K.M., 2021. A novel implicit hybrid machine learning model and its application for reinforcement learning. Comput. Chem. Eng. 155, 107496. https://doi.org/10.1016/j.compchemeng.2021.107496

Morales, J.L., Nocedal, J., 2011. Remark on "algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization." ACM Trans. Math. Softw. 38, 2–5. https://doi.org/10.1145/2049662.2049669

Narayanan, H., Luna, M., Sokolov, M., Butté, A., Morbidelli, M., 2022. Hybrid Models Based on Machine Learning and an Increasing Degree of Process Knowledge: Application to Cell Culture Processes. Ind. Eng. Chem. Res. 61, 8658–8672. https://doi.org/10.1021/acs.iecr.1c04507

Pantelides, C.C., Renfro, J.G., 2013. The online use of first-principles models in process operations: Review, current status and future needs. Comput. Chem. Eng. 51, 136–148. https://doi.org/10.1016/j.compchemeng.2012.07.008

Proctor, L., 2014. Method for the preparation of Dazoalkanes. US 2015/0038687 A1.

Psichogios, D.C., Ungar, L.H., 1992. A hybrid neural network-first principles approach to process modeling. AIChE J. 38, 1499–1511. https://doi.org/10.1002/aic.690381003

Riadi, I., 2014. Cognitive Ant colony optimization: A new framework in swarm intelligence. University of Salford.

Rojnuckarin, A., Floudas, C.A., Rabitz, H., Yetter, R.A., 1993. Optimal control of a plug flow reactor with a complex reaction mechanism. J. Phys. Chem. 97, 11689–11695. https://doi.org/10.1021/j100147a023

Sansana, J., Joswiak, M.N., Castillo, I., Wang, Z., Rendall, R., Chiang, L.H., Reis, M.S., 2021. Recent trends on hybrid modeling for Industry 4.0. Comput. Chem. Eng. 151, 107365. https://doi.org/10.1016/j.compchemeng.2021.107365

Schlueter, M., Egea, J.A., Antelo, L.T., Alonso, A.A., Banga, J.R., 2009. An Extended Ant Colony Optimization Algorithm for Integrated Process and Control System Design. Ind. Eng. Chem. Res. 48, 6723–6738.

Schwaab, M., Biscaia, E.C., Monteiro, J.L., Pinto, J.C., 2008. Nonlinear parameter estimation through particle swarm optimization. Chem. Eng. Sci. 63, 1542–1552. https://doi.org/10.1016/j.ces.2007.11.024

Shiva, M., Atashi, H., Farshchi Tabrizi, F., Mirzaei, A.A., Zare, A., 2013. The application of hybrid DOE/ANN methodology in lumped kinetic modeling of Fischer-Tropsch reaction. Fuel Process. Technol. 106, 631–640. https://doi.org/10.1016/j.fuproc.2012.09.056

Smets, I.Y., Dochain, D., Impe, J.F. Van, 2001. Steady-State Exothermic Plug Flow Reactor . Part I : Bang-Bang Control, in: 2001 European Control Conference (ECC). IEEE, Porto, Portugal, pp. 2540–2545. https://doi.org/10.23919/ECC.2001.7076310

Stoessel, F., 1997. Applications of reaction calorimetry in chemical engineering. J. Therm. Anal. 49, 1677–1688. https://doi.org/10.1007/bf01983728

von Stosch, M., Oliveira, R., Peres, J., Feyo de Azevedo, S., 2014. Hybrid semi-parametric modeling in process systems engineering: past, present and future. Comput. Chem. Eng. 60, 86–101.

Xie, Q., Liu, H., Bo, D., He, C., Pan, M., 2018. Data-driven Modeling and Optimization of Complex Chemical Processes Using a Novel HDMR Methodology. Comput. Aided Chem. Eng. 44, 835–840. https://doi.org/10.1016/B978-0-444-64241-7.50134-8

Zendehboudi, S., Rezaei, N., Lohi, A., 2018. Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review. Appl. Energy 228, 2539–2566. https://doi.org/10.1016/j.apenergy.2018.06.051