Chemical Engineering & Processing: Process Intensification 215 (2025) 110376

Contents lists available at ScienceDirect

Chemical Engineering and Processing - Process Intensification

journal homepage: www.elsevier.com/locate/cep

Prior knowledge-based multi-round multi-objective Bayesian optimization: continuous flow synthesis and scale-up of *O*-methylisourea

Jiapeng Guo [a], Kejie Chai [a,b], Guihua Luo [a], Weike Su [a,*], An Su [a,c,*]

[a] Key Laboratory of Pharmaceutical Engineering of Zhejiang Province, National Engineering Research Center for Process Development of Active Pharmaceutical Ingredients, Collaborative Innovation Center of Yangtze River Delta Region Green Pharmaceuticals, Zhejiang University of Technology, Hangzhou, 310014, China
[b] College of Material, Chemistry and Chemical Engineering, Hangzhou Normal University, Hangzhou, Zhejiang, 311121, China
[c] Zhejiang Key Laboratory of Green Manufacturing Technology for Chemical Drugs, College of Pharmaceutical Science, Zhejiang University of Technology, Hangzhou, Zhejiang, 310014, China

ARTICLE INFO

Keywords:
Bayesian optimization
Machine learning
Transfer learning
Continuous flow
O-methylisourea
Scale-up

ABSTRACT

Optimizing and scaling up chemical reactions is a critical step in transitioning from laboratory to industrial production, often involving trade-offs between multiple objectives such as production efficiency and cost. Meanwhile, the traditional synthesis process for the key pharmaceutical intermediate *O*-methylisourea no longer meets the demands of modern smart and green production, requiring urgent improvement. This study demonstrates the use of multi-objective Bayesian optimization (MOBO) for the continuous flow synthesis of *O*-methylisourea. By leveraging historical reaction data, we accelerated optimization as parameter ranges changed, enabling rapid inference and adjustment. Using a scaled-up continuous flow system equipped with MOBO and transfer learning capabilities, we identified optimal conditions along the Pareto front, achieving a production rate of up to 52.2 g/h and an E-factor as low as 0.557 during the third round of optimization, while maintaining a yield of approximately 75 %. These results highlight the scalability and efficiency of Bayesian optimization in accelerating reaction optimization and facilitating the transition from laboratory to industrial-scale production.

1. Introduction

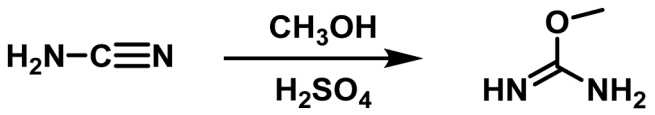
O-Methylisourea is a crucial pharmaceutical intermediate widely used in the synthesis of fluorouracil-based antitumor drugs, their derivatives, as well as pesticides, herbicides, polymers, and antimicrobial agents [1,2]. It is typically produced via an acid-catalyzed addition reaction between cyanamide and methanol in the presence of concentrated sulfuric acid (Scheme 1) [3]. However, the batch process for this reaction poses significant safety risks due to the toxicity of cyanamide and the use of concentrated sulfuric acid. Additionally, the highly exothermic nature of the reaction and the instability of cyanamide require strict temperature control at low levels in batch reactors, which limits the reaction rate and overall efficiency. In light of the growing emphasis on green and smart manufacturing in the chemical industry [4–6], the traditional process, characterized by safety hazards, high energy consumption, and low efficiency, urgently requires improvement.

Compared with traditional batch reactions [7], continuous-flow reactions offer significant advantages [8–14], providing a precise and reliable platform for reaction optimization [15,16]. This enables the use of methods such as one-factor-at-a-time (OFAT), design of experiments (DoE), and kinetic modeling [17]. However, these approaches often rely on the chemist's intuition [18–21] and may struggle to optimize multiple competing objectives simultaneously [22,23]. The development of multi-objective Bayesian optimization (MOBO) has addressed these challenges by efficiently balancing conflicting objectives, as represented by the Pareto front [24–28]. For example, Kappa et al. used MOBO to optimize the Buchwald-Hartwig amination reaction for olanzapine synthesis, achieving a 91 % yield and a maximum space-time yield of 1.13 kg/(L*h) in just 12 h [29]. Similarly, in our previous work, MOBO successfully optimized heterogeneous catalytic reactions for pharmaceutical intermediates, demonstrating its efficiency and effectiveness [30–32].

Traditional reaction optimization methods require significant resources when conditions change or scale up, whereas Bayesian optimization (BO) offers scalability by leveraging prior knowledge for rapid optimization. For example, Lapkin et al. used multitask BO to optimize a palladium-catalyzed C—H activation reaction and successfully

* Corresponding authors.
E-mail addresses: pharmlab@zjut.edu.cn (W. Su), ansu@zjut.edu.cn (A. Su).

https://doi.org/10.1016/j.cep.2025.110376
Received 6 January 2025; Received in revised form 2 May 2025; Accepted 23 May 2025
Available online 24 May 2025



Scheme 1. Synthesis of *O*-methylisourea from cyanamide.

transferred the results to synthesize an NK1 receptor antagonist with 82 % yield in just ten iterations [33]. MOBO further enhances versatility by optimizing multiple objectives simultaneously and expanding parameter ranges. Dunlap et al. used MOBO to optimize butylpyridinium bromide synthesis, generating a Pareto front and expanding the optimization space by increasing temperature limits [34]. Similarly, Slattery et al. improved yields in scalable, flow-based photochemical processes using MOBO [35]. Despite these advances, studies on extending MOBO results to new tasks or scale-up remain limited, highlighting the need for further research.

In this study, we optimized and scaled up the synthesis of *O*-methylisourea in a continuous flow system using MOBO. The first round of optimization established a preliminary Pareto front, followed by a second round with an expanded parameter range to further extend the frontier. By equipping the MOBO model with transfer learning capabilities [36], we enabled continued optimization in a scaled-up continuous flow system. Leveraging optimization data from the microreactor system, the method demonstrated its ability to rapidly optimize, expand the Pareto front, and effectively apply transfer learning, offering valuable guidance for reaction scale-up.

2. Results and discussion

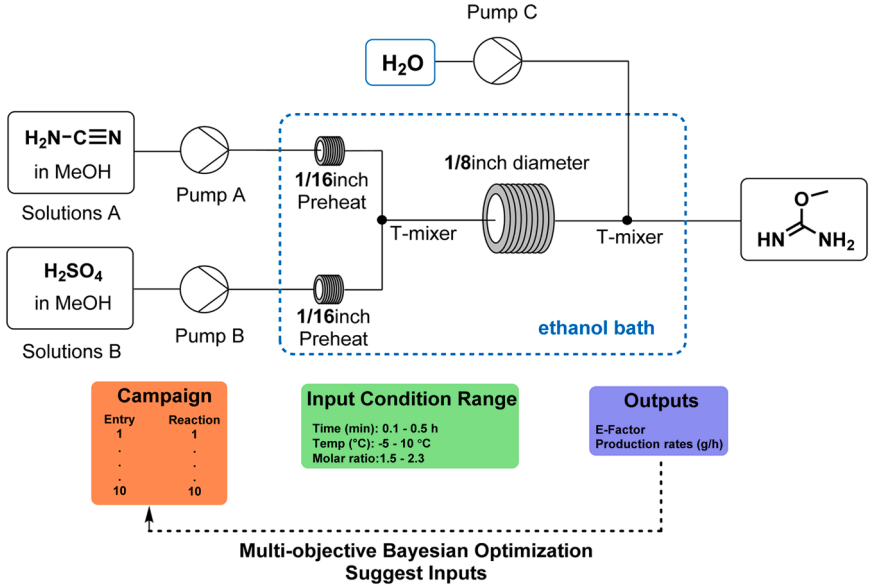
2.1. The first round of multi-objective Bayesian optimization

Multi-objective Bayesian optimizations (MOBO) were performed using the FlowBO platform developed by our group [30–32] to maximize the production rate and minimize the environmental factor (E-factor) (Scheme 2). The ranges for reaction parameters—temperature, residence time, and molar ratio—were initially defined based on preliminary investigations. If the parameter range is too narrow, optimization may overlook valuable regions in the reaction space. If it is too broad, the optimization model requires more experiments and more raw materials, increasing resource consumption. To conserve materials and improve efficiency, we adopted a stepwise approach: starting with a focused range for each parameter, and later expanding the ranges based on early results. This iterative strategy allows the optimization process to proceed efficiently while avoiding unnecessary waste. For our system, the initial temperature range was set to −5–10 °C to prevent cyanamide decomposition. The residence time was limited to 0.1–0.5 h, and the molar ratio of cyanamide to sulfuric acid was set between 1.5–2.3:1. In addition, due to cyanamide's limited solubility in methanol, its concentration was fixed at 0.6 mol/L.

The flowchart of MOBO is shown in Fig. 1. Initially, 10 experimental parameter sets were sampled using Latin Hypercube Sampling [37] (LHS) to efficiently cover the search space. In the iterative phase, a Gaussian process (GP) surrogate model was trained using the LHS data and corresponding experimental results. The acquisition function, *q* noisy expected hypervolume improvement (*q*NEHVI), was then used to identify the Pareto front and recommend the next experimental parameters. *q*NEHVI was chosen for its superior efficiency and ability to handle noise compared to other methods (e.g., *q*EHVI and *q*NParEGO) [25,27,32,38]. The algorithm selected a batch of experiments that maximized expected improvement over the current Pareto front, and the results were fed back to update the GP model. This process was repeated until the predetermined number of experiments was completed.

Fig. 2a shows the parameter distributions and experimental results for the 10 initial conditions selected using LHS, which adequately cover the reaction space and provide a solid foundation for the subsequent optimization experiments. After 10 optimization experiments, the results of the first round of MOBO were obtained (Fig. 2b). An initial Pareto front (red dashed line) was formed, balancing the trade-off between E-factor and production rate. The Pareto front demonstrates that the production rate can be increased from 3.48 g/h to 14.9 g/h at the cost of raising the E-factor from 0.459 to 1.16. Detailed parameters and results are provided in Table S1.

Fig. 2c provides an intuitive parallel coordinate visualization illustrating how each experimental variable—reaction time, temperature, and molar ratio—simultaneously impacts both production rate and E-factor. This plot allows clear identification of trends and trade-offs: for example, shorter reaction times generally lead to increased production rates but are often associated with higher E-factors, indicating less efficient reagent usage. Similarly, higher reaction temperatures and carefully tuned sulfuric acid ratios can be seen to influence both optimization objectives at the same time. Additional details on the distributions of these experimental variables are provided in the supplementary figures (Fig. S1).



Scheme 2. The workflow of continuous flow synthesis with Bayesian optimization.

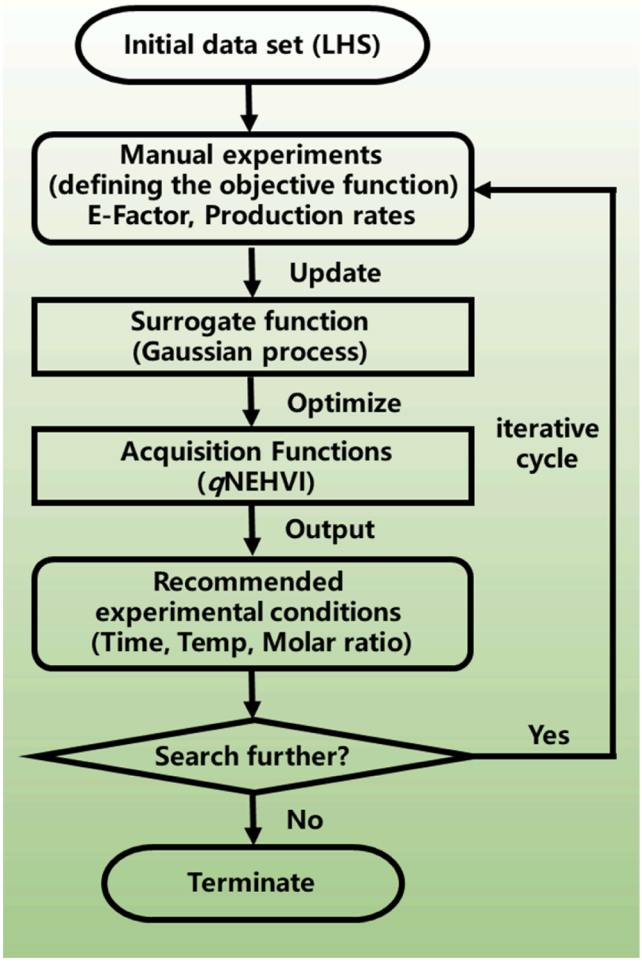


Fig. 1. The flowchart of MOBO in this study.

2.2. Extending the reaction space for a second round of optimization

In the first round of optimization experiments, reaction time and sulfuric acid molar ratio were found to significantly impact the production rate. To further explore the reaction space and extend the Pareto front, the parameter ranges were expanded. During our initial investigation, we observed that the material remains more stable at lower temperatures, and that the reaction conversion drops when the temperature exceeds 10 °C. Therefore, we kept the upper temperature limit at 10 °C to maintain product stability and conversion. By contrast, the reaction time range was increased from 0.1–0.5 h to 0.1–3 h, and the molar ratio from 1.5–2.3 to 0.5–2.3, aiming to achieve a lower E-factor and a higher production rate.

The Pareto front expands as the ranges for reaction time and mole ratios increase, resulting higher production rates and E-factors (Fig. 3a). In the second round of 10 MOBO iterations, equipped with the prior knowledge from the first round, the E-factor reached a minimum of 0.381 in the 17th experiment, while the production rate peaked at 17.5 g/h in the 19th experiment. However, production rates above 15 g/h resulted in E-factors exceeding 1.4 due to insufficient reaction time and excessive sulfuric acid using, highlighting the trade-off between production rate and E-factor.

The second round of optimization was performed with expanded parameter ranges compared to the first round, leading to a larger feasible search space and a corresponding increase in hypervolume. Therefore, the observed rise in hypervolume between rounds primarily illustrates the effect of expanding the parameter space, rather than simply the efficiency of the optimization algorithm. The hypervolume, which quantifies the space covered by the Pareto front, was calculated after each iteration to visualize the progression of trade-off improvements (Fig. 3b). During the first round, hypervolume increased from 0 to 4.99 as the Pareto front developed, but plateaued as the parameter space constraints limited further outcomes. In contrast, the second round's broader parameter space enabled additional exploration, resulting in substantial hypervolume growth that stabilized as the extended Pareto front was populated. This pattern is consistent with previous multi-stage optimization studies, such as the work by Baldwin et al. on pyridinium salt synthesis [34], where modifications to the studied parameter range were similarly tracked by changes in hypervolume as the optimization process progressed. The detailed experimental conditions, results, and hypervolume data are summarized in Tables S2 and S3.

2.3. MOBO in scale-up devices

2.3.1. Addressing mass transfer limitations in scale-up devices

Following two rounds of optimization in a continuous flow microreactor system, a scaled-up reaction device was constructed (Scheme 3) to evaluate its potential for industrial production and achieve higher production capacity. However, increasing the reactor tube diameter from 1/16 inch to 1/8 inch reduced mass transfer efficiency, which limits the unit's capacity [39]. To address this, a simple and cost-effective static mixer was designed to enhance mixing efficiency. The impact of flow rate on reaction conversion was then assessed under two different mixing conditions (Fig. 4a and 4c).

The evaluation was conducted at a reaction temperature of 10 °C. Using only the T-mixer, as shown in Fig. 4a, the effect of flow rate on conversion was assessed (Fig. 4c). Even at flow rate of 686 mL/h, the conversion did not stabilize at a plateau, indicating unresolved mass transfer limitations. In contrast, with our homemade static mixer (a 1/4-inch mixing coil filled with SiO_2 beads, Fig. 4b), the conversion plateaued once the total flow rate exceeded 201 mL/h (Fig. 4d), demonstrating the elimination of mass-transfer limitations. This improvement in mixing efficiency is attributed to the mixer's design, which incorporates double reverse rotating vortices [40,41], generated by the helical coils and turbulence induced by obstacles within the channel [42, 43].

2.3.2. MOBO based on prior knowledge

When the reaction was scaled up, the optimization task changed, and the previous MOBO model could not be directly reused without adaptation. To address this, we enhanced the optimization model to incorporate learning from prior data, thus enabling transfer of relevant knowledge from small-scale experiments to the scaled-up system. This approach, similar to transfer learning in BO [33,44], allows the model to selectively extract and use previous results, thereby accelerating optimization on the larger-scale system. This methodology is closer to common industrial workflows, where reactions are first optimized at small scale and then further refined during scale-up. Our model required only three new experimental runs to initialize the scaled-up optimization, significantly reducing the experimental effort required. Using this enhanced model and the prior optimization datasets, a third round of optimization was conducted (Scheme 4), incorporating the homemade static mixer. The time ranges and molar ratios maintained the expanded parameter ranges in the second round of optimization, i.e., a time range of 0.1–3 h, a molar ratio of cyanamide to H_2SO_4 of 0.5–2.3, and a temperature range of −5–10 °C.

Fig. 5a shows the Pareto front for the third round of optimization, which has shifted significantly from the previous two rounds due to the increased production rate achieved by scaling up the device. The new Pareto front shifted away from the original one, with a minimum E-factor of 0.557 and a production rate of 39.9 g/h, and a maximum production rate of 52.2 g/h with an E-factor of 0.686. This shift reflects the Bayesian optimization model's ability to learn from prior rounds, focusing on low E-factor, high production rate conditions while avoiding

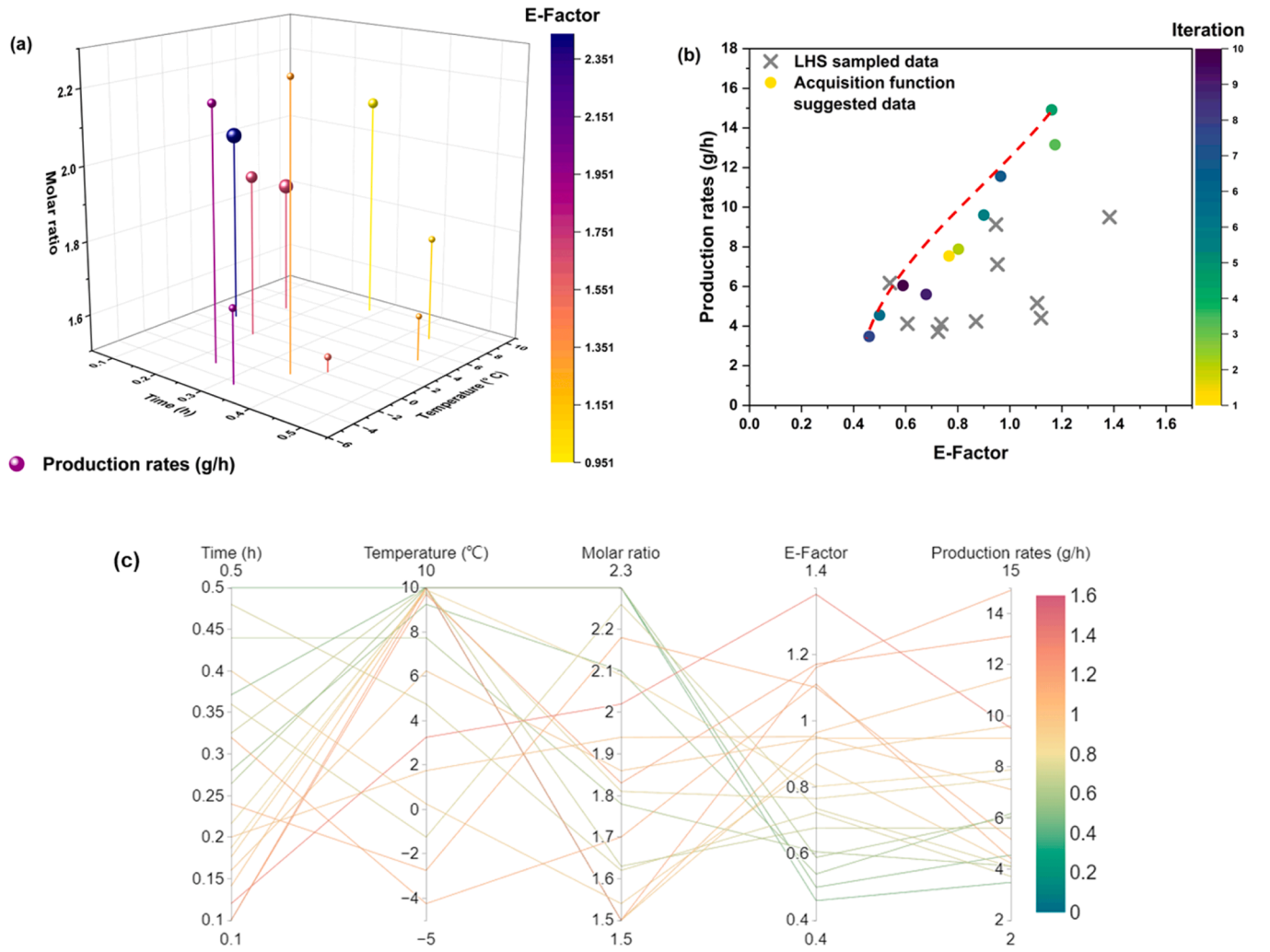


Fig. 2. (a) The first round of optimization: LHS sampling stage. (b) Results of the first round of optimization (red dotted line denotes the Pareto front) (c) Parallel coordinate plot showing the integrated effects of reaction time, temperature, and molar ratio on both production rate and E-factor.

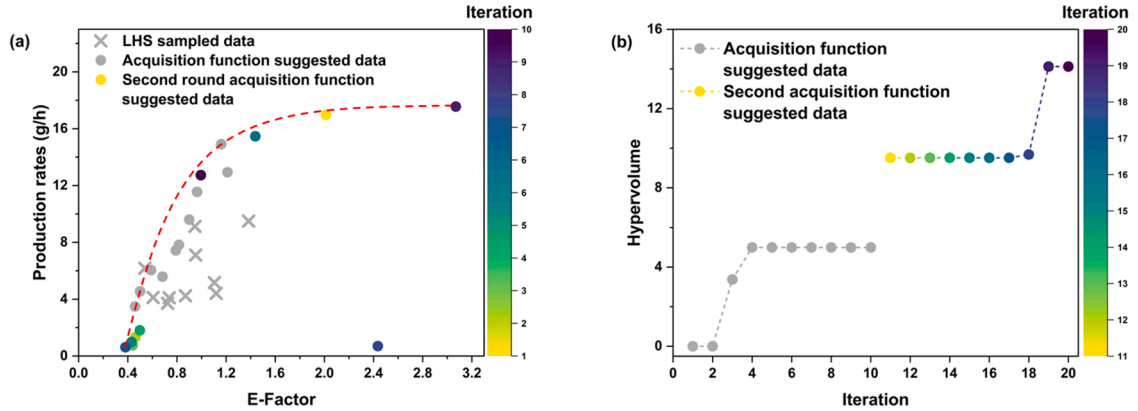
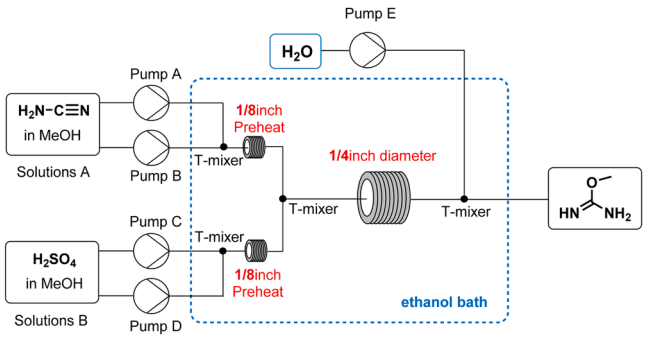


Fig. 3. (a) The second round of optimization results. (The red dotted line is the Pareto front.) (b) Expansion of the hypervolume of all solutions to the Pareto front. All the hypervolume were scaled according to Eq. (5) in the section of Materials and Methods.

high E-factor, high production rate or low E-factor, low production rate scenarios.

As shown in Fig. 5b-d, reaction time was concentrated between 0.260–0.326 h, favoring shorter time for higher production rate. The reaction temperature consistently remained at its upper limit to accelerate the reaction, while the molar ratio ranged from 1.82:1–2.30:1, minimizing sulfuric acid usage to reduce the E-factor. Although the E-factor was slightly higher than the lowest value in the second round (remaining between 0.5–0.6, Fig. 5e), the production rate improved significantly, consistently exceeding about 40 g/h (Fig. 5f).

Compared to optimized result in the first-round optimization, the final optimized results in the scaled-up device achieved a more than



Scheme 3. Continuous flow synthesis setup for scale-up.

tenfold increase in production rate, with only a slight E-factor increase of about 0.2, which remained at a low level. The high consistency of reaction parameters and results in this round indicates a well-balanced trade-off between the E-factor and production rate, with an average yield of 75 %. Additionally, the optimized process features a significantly shorter reaction time and is intrinsically safer than the original process [45]. Detailed parameters and results are provided in Table S4.

3. Conclusion

In this work, we demonstrated the effectiveness of multiple rounds of MOBO in the continuous flow synthesis of *O*-methylisourea, both at the small-scale and in a scaled-up device. By iteratively refining the compromise between higher production rates and lower E-factors, we achieved an ever-improving Pareto front, culminating in optimized scaled-up production conditions with a maximum production rate of 52.2 g/h, a minimum E-factor of 0.557, and a maintained yield of approximately 75 %.

The MOBO model proved highly adaptable, responding efficiently to changes in parameter ranges and advancing the Pareto front into previously inaccessible regions. Moreover, the incorporation of transfer learning capabilities allowed the model to leverage historical optimization data, enabling seamless transitions from laboratory-scale to industrial-scale optimization. This approach significantly reduced the time required for reaction optimization while achieving a completely new Pareto front tailored to the scaled-up system. The application of this method provides a powerful and efficient framework for optimizing complex chemical processes and holds great promise for guiding the large-scale production of pharmaceutical intermediates and other fine chemicals.

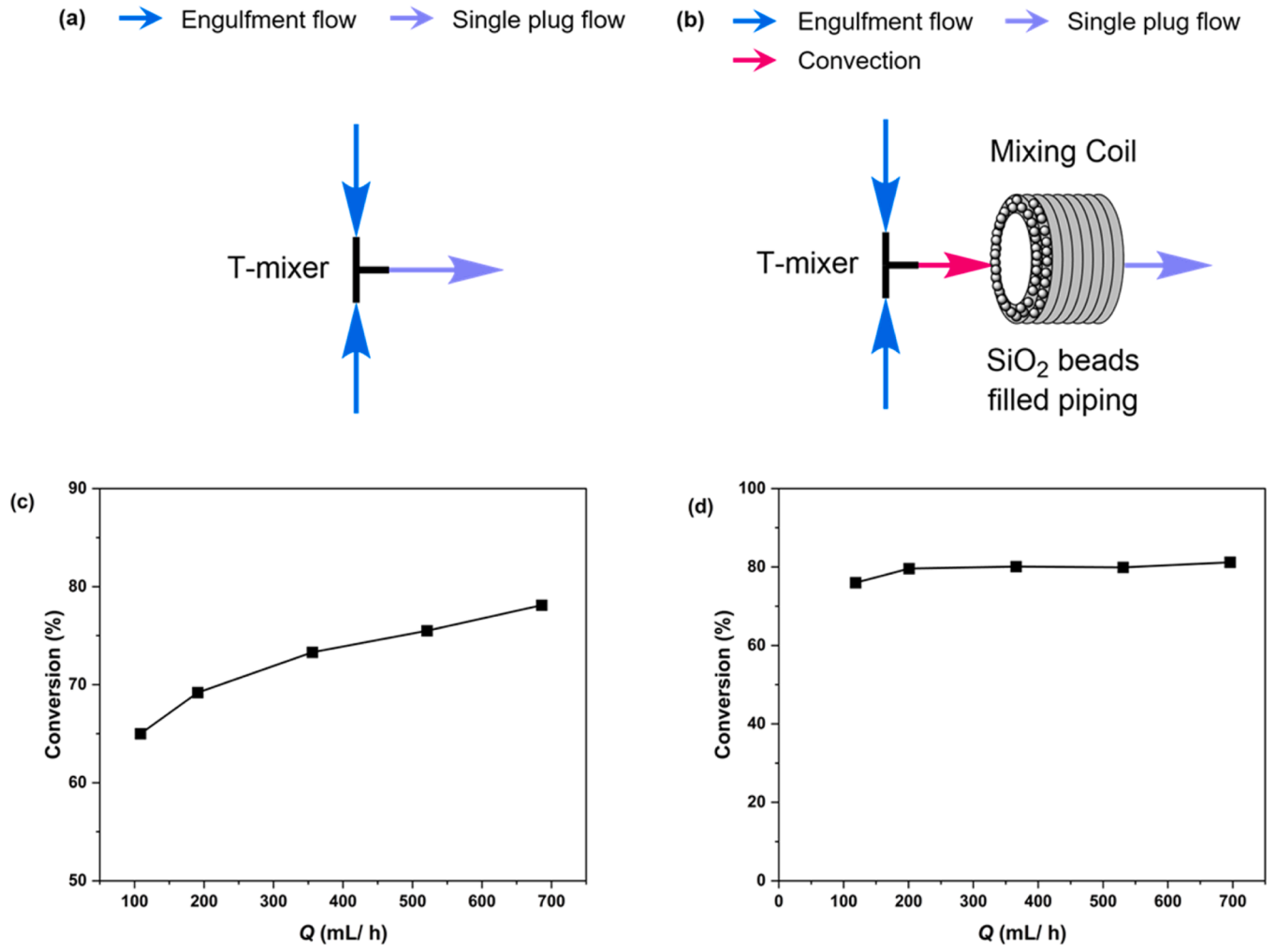
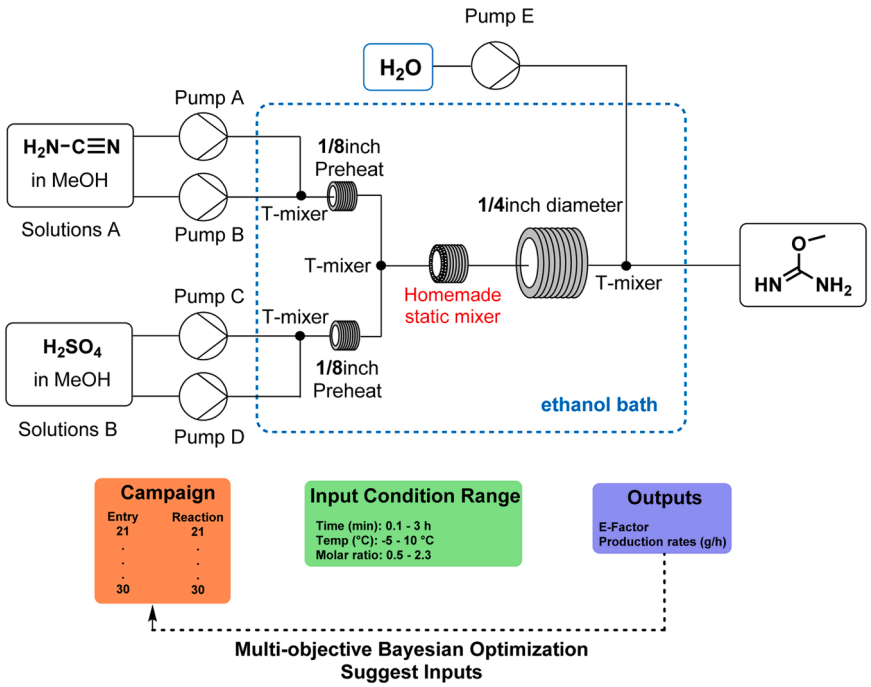


Fig. 4. Two mixing setups: (a) a T-mixer and (b) a T-mixer combined with a homemade static mixer, and the effect of the two mixing setups on the mixing process; (c) the T-mixer and (d) the T-mixer plus the homemade static mixer effect of flow rate on conversion. Reaction conditions: residence time t = 0.1 h, reaction temperature T = 10 °C, molar ratio of cyanamide to H_2SO_4 = 0.5:1, Initial concentration of reactants $c_{cyanamide}$ = 0.6 mol/L.



Scheme 4. Continuous flow synthesis setup and Bayesian optimization workflow for scale-up.

4. Materials and methods

4.1. Chemicals

Cyanamide (95 %) purchased from Shanghai Macklin Biochemical Technology Co., Ltd.; Methanol (99.9 %) purchased from Sinopharm Chemical Reagent Co., Ltd.; H_2SO_4 (98 %, Sinopharm Chemical Reagent Co., Ltd.); pure water (AR, Hangzhou Wahaha Group Co., Ltd.); All reagents were used without further purification.

Solution A: Cyanamide (0.02 mol, 2.382 g) was dissolved in Methanol (50 mL).

Solution B: H_2SO_4 (concentration adjusted to conditions) was dissolved in stirred under ice bath conditions Methanol (50 mL).

4.2. Continuous flow microreactor system

The continuous flow microreactor system is shown in Scheme 2. Solutions A and B were stored in two separate syringes (50 mL, Shanghai Kindly Medical Instruments Co., Ltd.) and pumped by two syringe pumps (Pump A, Pump B, ZD-50C6, Suzhou Zede Medical Instrument Co., Ltd.) into the infusion tubing (PTFE, 1/8inch diameter, Wuxi Hongxin Special Material Technology Co.). The reactants passed through an adapter (PEEK, 1/8inch to 1/16inch, Hangzhou Weimipai Technology Co.) into a sufficiently long (1 m) disk-shaped PTFE capillary tube (PTFE, 1/16inch diameter, Wuxi Hongxin Special Material Technology Co.) to be preheated to the reaction temperature, and mixed in a T-mixer (PEEK, 1/16inch to 1/8inch, Hangzhou Weimipai Technology Co.). The reaction coil (PTFE, 1/8inch diameter) was connected directly to the outlet of the T- mixer. The addition reaction taken placed in the reaction coil, and the residence time can be precisely controlled by varying the flow rate of the reaction mixture or the length of the reaction coil. All preheat tubes, T-mixer, and reaction coils were immersed in the same water bath to maintain a constant temperature. Finally, when the required residence time was reached, the reaction was terminated by pumping an excess of purified water through a syringe pump (Pump C, ZD-50C6, Suzhou Zede Medical Instrument Co., Ltd.) into a second T-mixer.

4.3. Scale-up of continuous flow system

The Scale-up of continuous flow system is shown in Scheme 3 and Scheme 4 (Scheme 4 is identical to Scheme 3 only with the addition of a homemade static mixer.). Solutions A and B were stored in four separate syringes (50 mL, Shanghai Kindly Medical Instruments Co., Ltd.) and pumped by four syringe pumps (Pump A, Pump B, and Pump C, Pump D, ZD-50C6, Suzhou Zede Medical Instrument Co., Ltd.) into the infusion tubing (PTFE, 1/8inch diameter, Wuxi Hongxin Special Material Technology Co.). The reactants into a sufficiently long (1 m) disk-shaped stainless steel metal tube (SS316L, 1/8inch diameter, Beijing Xiongchuan Technology Co., Ltd.) to be preheated to the reaction temperature, and mixed in a T-mixer (SS316L, 1/8inch to 1/4inch, Beijing Xiongchuan Technology Co., Ltd.). The reaction solution enters a homemade static mixer (SiO_2 beads, 3 mm diameter; piping, 1/4-inch diameter, 30 cm length, Wuxi Hongxin Special Material Technology Co.) for further mixing. The reaction coil (PTFE, 1/4inch diameter, Wuxi Hongxin Special Material Technology Co.) was connected directly to the outlet of the T-mixer. The reaction taken placed in the reaction coil, and the residence time can be precisely controlled by varying the flow rate of the reaction mixture or the length of the reaction coil. All preheat tubes, T-mixer, and reaction coils were immersed in the same ethanol bath to maintain a constant temperature. Finally, when the required residence time was reached, the reaction was terminated by pumping an excess of purified water through a high-pressure PTFE pump (Pump E, JJRZ-10004F, Hangzhou JingJin Technology Co., Ltd.) into a second T-mixer.

4.4. Sample analysis

The reaction solutions were collected for quenching at appropriate times and analyzed by high-performance liquid chromatography (HPLC, ThermoFisher Ulcel3000). The conversion rate of the samples was obtained by constructing the regression equation of the standard curve by the external standard method. HPLC detection conditions C18 column (10 μm, 4.6 × 250 mm, Welch Materials (Shanghai, China), USA), the mobile phase was 15 % acetonitrile, 85 % ultrapure water and 600 mmol/L phosphoric acid at a flow rate of 0.8 mL/min, and the detection wavelength was 200 nm. The E-Factor was calculated by the following

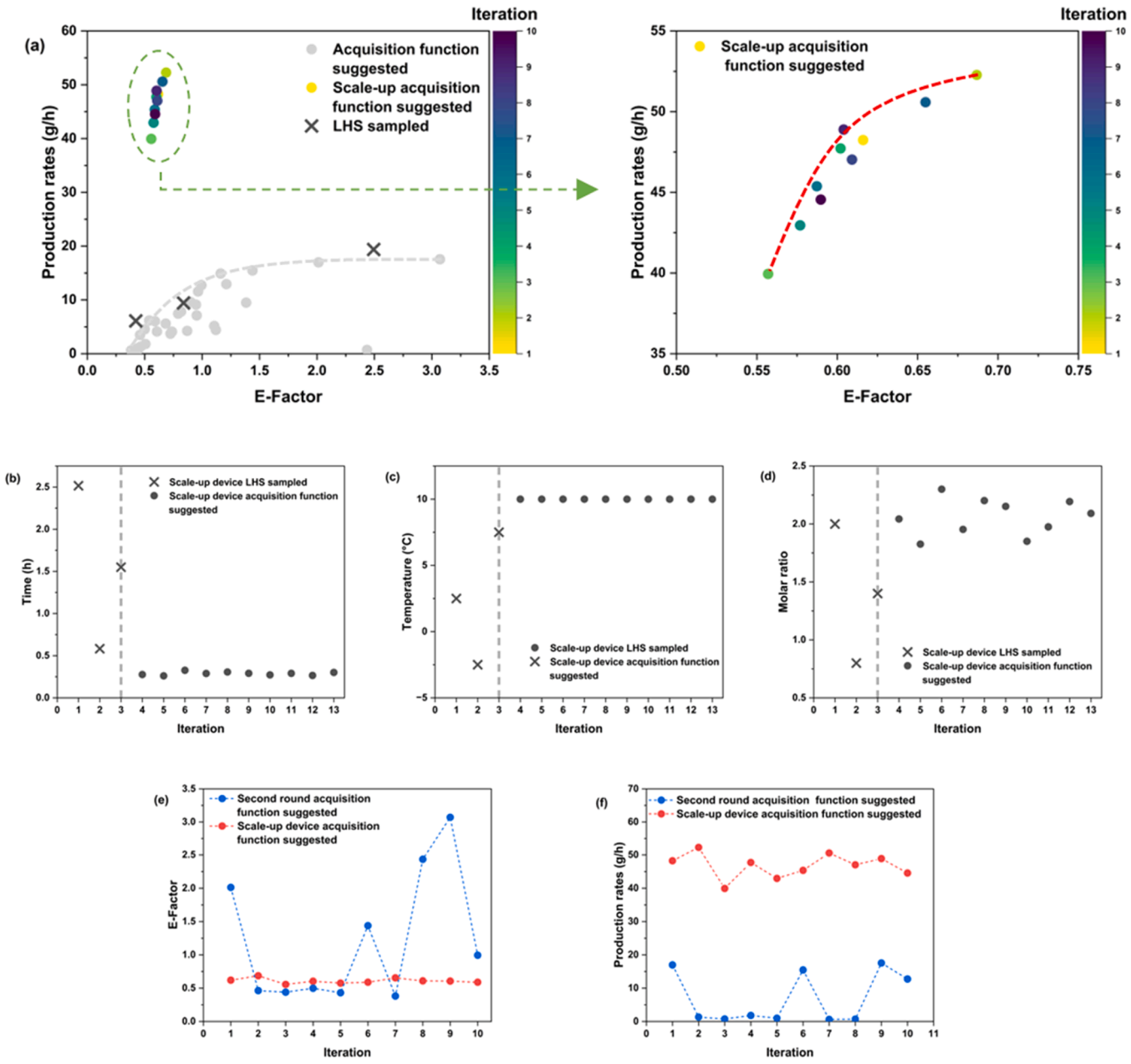


Fig. 5. (a) The third round of optimization LHS sampling stage and optimization results. (The data in the green dashed box have been scaled for easier viewing, the red dotted line and gray dotted line are the Pareto front.) (b-d) The third round of optimization exploration of three decision variables. (e) E-factor results for the second and third optimization stages. (f) Production rate results for the second and third optimization stages.

equation [46]:

$$E-Factor = \left(\frac{m_{waste}}{m_{product}}\right) \quad (1)$$

where m_{waste} is the mass of the wasted raw material and $m_{product}$ is the mass of the product.

Separation and purification are important aspects of evaluating reaction sustainability, and the waste produced during these steps can significantly contribute to the overall E-factor, particularly in processes with lower yields. However, most post-processing operations are time-consuming and would slow down the feedback loop needed for efficient optimization. Therefore, in this study, the E-factor calculation only includes the mass of wasted raw materials, in order to maintain a rapid and efficient optimization process.

The Production rate was calculated by the following equation [34]:

$$Production\ rate = \frac{m_{product}}{t_{res}} \quad (2)$$

where t_{res} is the residence time of reaction.

Samples were tested three times under the same conditions and averaged to minimize errors.

4.5. Bayesian optimization and transfer learning

The BO model is composed of three core components: initial sampling, a surrogate model, and an acquisition function [37,47]. For initial sampling, LHS is employed to prevent the over-aggregation of data that can occur with simple random sampling. LHS divides the sampling space into distinct strata based on specific characteristics or rules, and samples are then drawn independently and randomly from each stratum.

Notably, LHS requires fewer samples to achieve the same level of accuracy compared to traditional methods such as One-Factor-At-a-Time (OFAT) or Design of Experiments (DoE), thereby reducing computational complexity.

Bayesian Optimization (BO) is an uncertainty-guided response surface method, where the performance of the surrogate model reflects the optimizer's prediction accuracy. The effectiveness of a surrogate model is demonstrated only when its expectation and variance estimates closely approximate the true response surface [48]. The Gaussian Process (GP), an infinite-dimensional extended function distribution, is commonly used as a surrogate model [44]. GP enables the construction of joint probability distributions for variables, allowing the estimation of both the mean and variance of predicted data based on the available data. The Matérn class is a widely used class of covariance functions in GP, as described by Eq. (3) [49]:

$$M_{Matèrn}(x,y) = \frac{2^{1-\nu}\sigma^2}{\Gamma(\nu)}\left(\sqrt{2\nu}\parallel x-y\parallel\right)^{\nu} K_{\nu}\left(\sqrt{2\nu}\parallel x-y\parallel\right) \tag{3}$$

Where, $\sigma^2 > 0, \nu > 0; \nu, \sigma^2, \Gamma(\nu), \parallel x-y \parallel$ and K_ν represent non-negative parameters, output variance, Gamma function, distance between two points and Bessel function, respectively.

In addition, the acquisition function is crucial for the desired optimization performance. Among others, probabilistic improvements, expected improvements, and confidence upper bounds are often used to tune the hyperparameters. The q noisy expected hypervolume improvement (qNEHVI) function was used as the acquisition function, *via* Eq. (4) [50]:

$$\alpha_{qNEHVI}(\chi_{cand}|\mathscr{P}) = \frac{1}{\widetilde{N}}\sum_{t=1}^{\widetilde{N}} HVI(\widetilde{f}_t(\chi_{cand}|\mathscr{P})) \tag{4}$$

Where, $\widetilde{N}$, HVI, $\widetilde{f}_t$, χ_{cand} and $\mathscr{P}$ stand for the number of samples, hypervolume improvement, Black-box objective function, candidate sample, and Pareto front, respectively.

In this experiment, the E-factor of the product and the production rate were selected as the optimization objectives for MOBO. Using MOBO, the Pareto front for the continuous flow synthesis of *O*-methylisourea was rapidly obtained to optimize the reaction process.

The reference point is a critical parameter in hypervolume calculation for multi-objective optimization. In this study, following the methodology reported by Baldwin et al. in EDBO+ [34], we dynamically define the reference point as the vector of minimum observed values for each objective (e.g., yield and space-time yield) obtained in the experiments. Because the optimization is performed as a black-box procedure without a fixed, predefined reference, this dynamic approach ensures the reference point always reflects the most current results, and is updated after each iteration accordingly.

The equation for hypervolume scaling is as follows:

$$\text{hypervolume} = \frac{H - x_1}{C} \tag{5}$$

where H is the hypervolume obtained from the computation, x_1 is the first optimized hypervolume in the first round, and C is a custom scaling factor.

For the scale-up, we implemented MOBO with transfer learning capabilities to enable the application of knowledge across different optimization campaigns. The model leverages neural processes (NPs) [51–53] to guide the search space targeted by the BO strategy. NPs are a class of neural models that define the conditional distribution of a function based on a set of observed measurements. The enhanced MOBO trains the NPs using data points from individual historical campaigns. During the optimization of a new target campaign, the NPs generate a distribution of functions for the target campaign based on the current set of collected measurements, allowing predictions about other points in the search space. These predictions are then used to adjust the acquisition function for the target campaign. This approach combines a meta-learning strategy [54] with Bayesian optimization, enabling the transfer of knowledge from source campaigns to the target campaign. Further details about the algorithm can be found in previous studies [40, 41].

CRediT authorship contribution statement

Jiapeng Guo: Writing – original draft, Visualization, Validation, Investigation, Formal analysis, Data curation, Conceptualization. **Kejie Chai:** Writing – review & editing, Investigation. **Guihua Luo:** Writing – review & editing, Software. **Weike Su:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization. **An Su:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the Joint Funds of the Zhejiang Provincial Natural Science Foundation of China under Grant No. LHDMZ23B060001, Zhejiang Province Science and Technology Plan Project under Grant No. 2022C01179, and the National Natural Science Foundation of China under Grant No. 22108252.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cep.2025.110376.

Data availability

No data was used for the research described in the article.

References

[1] Wang, M.W., .Jian feng; Ren, Ju; Si, Shuangxi; Wen, Jianhua The Invention Relates to a Process for the Preparation of O-Me Isourea Via Methylation of Di-Me Carbonate With Urea in the Presence of NaOH (or KOH). CN102432506, 2012.
[2] Thomas, C. Preparation of Carbocyclic Nucleoside Analogs As Antitumor agents. EP3747872, 2020.
[3] Elliott, A.J.; Walsh, D.A.; Morris, P.E. P.rocess for the Preparation of 9-deazaguanine derivatives. US20040254181, 2004.
[4] C. Wiles, P. Watts, Continuous flow reactors: a perspective, Green Chem. 14 (1) (2012) 38–54, https://doi.org/10.1039/C1GC16022B.
[5] C. Wiles, P. Watts, Continuous process technology: a tool for sustainable production, Green Chem. 16 (1) (2014) 55–62, https://doi.org/10.1039/C3GC41797B.
[6] P. Liu, H. Jin, Y. Chen, D. Wang, H. Yan, M. Wu, F. Zhao, W. Zhu, Process analytical technologies and self-optimization algorithms in automated pharmaceutical continuous manufacturing, Chin. Chem. Lett. 35 (3) (2024) 108877, https://doi.org/10.1016/j.cclet.2023.108877.
[7] K. Chai, Z. Ni, L. Chen, L. Chen, Y. Wen, Y. Huang, P. Zhang, W. Xu, Pd particles decorated 2D-MoSe2 nanomesh as a distinctive catalyst for semihydrogenation of alkynes, Chem. Eng. J. 497 (2024) 154752, https://doi.org/10.1016/j.cej.2024.154752.
[8] Hartman, R.L.; McMullen, J.P.; Jensen, K.F. D.eciding Whether To Go with the Flow: Evaluating the Merits of Flow Reactors for Synthesis. 2011, 50 (33), 7502–7519, https://doi.org/10.1002/anie.201004637.
[9] M.R. Chapman, M.H.T. Kwan, G. King, K.E. Jolley, M. Hussain, S. Hussain, I. E. Salama, C. González Niño, L.A. Thompson, M.E. Bayana, A.D. Clayton, B. N. Nguyen, N.J. Turner, N. Kapur, A.J. Blacker, S.imple and versatile laboratory scale CSTR for multiphasic continuous-flow chemistry and long residence times, Org. Process Res. Dev. 21 (9) (2017) 1294–1301, https://doi.org/10.1021/acs.oprd.7b00173.
[10] Gutmann, B.; Cantillo, D.; Kappe, C.O. C.ontinuous-Flow Technology—A Tool for the Safe Manufacturing of Active Pharmaceutical Ingredients. 2015, 54 (23), 6688–6728, doi:https://doi.org/10.1002/anie.201409318.

[11] Razzaq, T.; Glasnov, T.N.; Kappe, C.O. C.ontinuous-Flow Microreactor Chemistry under High-Temperature/Pressure Conditions. 2009, *2009* (9), 1321–1325, doi:https://doi.org/10.1002/ejoc.200900077.
[12] M. Köckinger, B. Wyler, C. Aellig, D.M. Roberge, C.A. Hone, C.O. Kappe, O. ptimization and scale-up of the continuous flow acetylation and nitration of 4-Fluoro-2-methoxyaniline to prepare a key building block of Osimertinib, Org. Process Res. Dev. 24 (10) (2020) 2217–2227, https://doi.org/10.1021/acs.oprd.0c00254.
[13] G.-N. Ahn, J.-H. Kang, H.-J. Lee, B.E. Park, M. Kwon, G.-S. Na, H. Kim, D.-H. Seo, D.-P. Kim, Exploring ultrafast flow chemistry by autonomous self-optimizing platform, Chem. Eng. J. 453 (2023) 139707, https://doi.org/10.1016/j.cej.2022.139707.
[14] G. Fu, L. Ni, D. Wei, J. Jiang, Z. Chen, Y. Pan, Scale-up and safety of toluene nitration in a meso-scale flow reactor, Process Saf. Environ. Prot. 160 (2022) 385–396, https://doi.org/10.1016/j.psep.2022.02.036.
[15] Ingham, R.J.; Battilocchio, C.; Fitzpatrick, D.E.; Sliwinski, E.; Hawkins, J.M.; Ley, S. V. A. Systems Approach towards an Intelligent and Self-Controlling Platform for Integrated Continuous Reaction Sequences. 2015, 54 (1), 144–148, doi:https://doi.org/10.1002/anie.201409356.
[16] J. Guo, W. Su, A. Su, Homogeneous continuous flow nitration of O-methylisouronium sulfate and its optimization by kinetic modeling, Beilstein J. Org. Chem. 20 (2024) 2408–2420, https://doi.org/10.3762/bjoc.20.205.
[17] C.J. Taylor, A. Pomberger, K.C. Felton, R. Grainger, M. Barecka, T.W. Chamberlain, R.A. Bourne, C.N. Johnson, A.A. Lapkin, A. brief introduction to chemical reaction optimization, Chem. Rev. 123 (6) (2023) 3089–3126, https://doi.org/10.1021/acs.chemrev.2c00798.
[18] C.T. Brain, J.T. Steer, A.n improved procedure for the synthesis of benzimidazoles, using palladium-catalyzed aryl-amination chemistry, J. Org. Chem. 68 (17) (2003) 6814–6816, https://doi.org/10.1021/jo034824L.
[19] J.B. Roque, Y. Kuroda, L.T. Göttemann, R. Sarpong, Deconstructive diversification of cyclic amines, Nature 564 (7735) (2018) 244–248, https://doi.org/10.1038/s41586-018-0700-3.
[20] H. Wang, H. Jung, F. Song, S. Zhu, Z. Bai, D. Chen, G. He, S. Chang, G. Chen, Nitrene-mediated intermolecular N–N coupling for efficient synthesis of hydrazides, Nat. Chem. 13 (4) (2021) 378–385, https://doi.org/10.1038/s41557-021-00650-0.
[21] I. Kim, H. Im, H. Lee, S. Hong, N-Heterocyclic carbene-catalyzed deaminative cross-coupling of aldehydes with Katritzky pyridinium salts, Chem. Sci. 11 (12) (2020) 3192–3197, https://doi.org/10.1039/D0SC00225A.
[22] S.M. Awerinde, A.M. Schweidtmann, A.A. Lapkin, T.he concept of selectivity control by simultaneous distribution of the oxygen feed and wall temperature in a microstructured reactor, Chem. Eng. J. 331 (2018) 765–776, https://doi.org/10.1016/j.cej.2017.09.030.
[23] Abtahi, B.; Tavakol, H. Choline chloride-urea deep eutectic solvent as an efficient media for the synthesis of propargylamines via organocuprate intermediate. 2020, 34 (11), e5895, doi:https://doi.org/10.1002/aoc.5895.
[24] R. Tachibana, K. Zhang, Z. Zou, S. Burgener, T.R. Ward, A. Customized bayesian algorithm to optimize enzyme-catalyzed reactions, ACS Sustain. Chem. Eng. 11 (33) (2023) 12336–12344, https://doi.org/10.1021/acssuschemeng.3c02402.
[25] Deb, K.; Sindhya, K.; Hakanen, J., Multi-Objective Optimization: Theory and Practice. 2016; pp 145–184.
[26] D.N. Jumbam, R.A. Skilton, A.J. Parrott, R.A. Bourne, M. Poliakoff, The Effect of Self-Optimisation Targets on the Methylation of Alcohols Using Dimethyl Carbonate in Supercritical CO2, J. Flow Chem. 2 (1) (2012) 24–27, https://doi.org/10.1556/jfchem.2012.00019.
[27] J. Guo, G. Luo, K. Chai, W. Su, A. Su, Continuous flow synthesis of N,O-Dimethyl-N′-nitroisourea monitored by inline fourier transform infrared spectroscopy: bayesian optimization and kinetic modeling, Ind. Eng. Chem. Res. 63 (23) (2024) 10162–10171, https://doi.org/10.1021/acs.iecr.4c01003.
[28] B.J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J.I.M. Alvarado, J.M. Janey, R.P. Adams, A.G. Doyle, B.ayesian reaction optimization as a tool for chemical synthesis, Nature 590 (7844) (2021) 89–96, https://doi.org/10.1038/s41586-021-03213-y.
[29] F. Wagner, P. Sagmeister, C.E. Jusner, T.G. Tampone, V. Manee, F.G. Buono, J. D. Williams, C.O. Kappe, A. slug flow platform with multiple process analytics facilitates flexible reaction optimization, Adv. Sci. 11 (13) (2024) 2308034, https://doi.org/10.1002/advs.202308034.
[30] T. Qi, G. Luo, H. Xue, F. Su, J. Chen, W. Su, K.-J. Wu, A. Su, Continuous heterogeneous synthesis of hexafluoroacetone and its machine learning-assisted optimization, J. Flow Chem. 13 (3) (2023) 337–346, https://doi.org/10.1007/s41981-023-00273-1.
[31] K. Chai, W. Xia, R. Shen, G. Luo, Y. Cheng, W. Su, A. Su, Optimization of heterogeneous continuous flow hydrogenation using FTIR inline analysis: a comparative study of multi-objective Bayesian optimization and kinetic modeling, Chem. Eng. Sci. 302 (2025) 120901, https://doi.org/10.1016/j.ces.2024.120901.
[32] G. Luo, X. Yang, W. Su, T. Qi, Q. Xu, A. Su, Optimizing telescoped heterogeneous catalysis with noise-resilient multi-objective Bayesian optimization, Chem. Eng. Sci. 298 (2024) 120434, https://doi.org/10.1016/j.ces.2024.120434.
[33] C.J. Taylor, K.C. Felton, D. Wigh, M.I. Jeraal, R. Grainger, G. Chessari, C. N. Johnson, A.A. Lapkin, A.ccelerated chemical reaction optimization using multi-task learning, ACS Cent. Sci. 9 (5) (2023) 957–968, https://doi.org/10.1021/acscentsci.3c00050.
[34] J.H. Dunlap, J.G. Ethier, A.A. Putnam-Neeb, S. Iyer, S.-X.L. Luo, H. Feng, J. A. Garrido Torres, A.G. Doyle, T.M. Swager, R.A. Vaia, P. Mirau, C.A. Crouse, L. A Baldwin, Continuous flow synthesis of pyridinium salts accelerated by multi-objective Bayesian optimization with active learning, Chem. Sci. 14 (30) (2023) 8061–8069, https://doi.org/10.1039/D3SC01303K.
[35] A. Slattery, Z. Wen, P. Tenblad, J. Sanjosé-Orduna, D. Pintossi, T. den Hartog, T. Noël, Automated self-optimization, intensification, and scale-up of photocatalysis in flow, Science 383 (6681) (2024) eadj1817, https://doi.org/10.1126/science.adj1817.
[36] K. Weiss, T.M. Khoshgoftaar, D. Wang, A survey of transfer learning, J. Big. Data 3 (1) (2016) 9, https://doi.org/10.1186/s40537-016-0043-6.
[37] A. Navid, S. Khalilarya, M. Abbasi, Diesel engine optimization with multi-objective performance characteristics by non-evolutionary Nelder-Mead algorithm: sobol sequence and Latin hypercube sampling methods comparison in DoE process, Fuel 228 (2018) 349–367, https://doi.org/10.1016/j.fuel.2018.04.142.
[38] B. Letham, B. Karrer, G. Ottoni, E. Bakshy, Constrained bayesian optimization with noisy experiments, Bayesian Anal 14 (2) (2019) 495–519.
[39] P. Watts, C. Wiles, Micro reactors, flow reactors and continuous flow synthesis, J. Chem. Res. 36 (2012) 181–193, https://doi.org/10.3184/174751912X13311365798808.
[40] W.R. Dean, J.M. Hurst, N.ote on the motion of fluid in a curved pipe, Mathematika 6 (1) (1959) 77–85, https://doi.org/10.1112/S0025579300001947.
[41] C. Jonas Bolinder, B. Sundén, Flow visualization and LDV measurements of laminar flow in a helical square duct with finite pitch, Exp. Therm. Fluid Sci. 11 (4) (1995) 348–363, https://doi.org/10.1016/0894-1777(95)00040-2.
[42] L. Tata Rao, S. Goel, S. Kumar Dubey, A. Javed, Performance investigation of T-shaped micromixer with different obstacles, J. Phys.: Conf. Ser. 1276 (1) (2019) 012003, https://doi.org/10.1088/1742-6596/1276/1/012003.
[43] Y. Bazargan-Lari, S. Movahed, M. Mashhoodi, Control of flow field, mass transfer and mixing enhancement in T-shaped microchannels, J. Mech. 33 (3) (2016) 387–394, https://doi.org/10.1017/jmech.2016.81. %J Journal of Mechanics.
[44] M. Feurer, B. Letham, F. Hutter, E. Bakshy, Practical transfer learning for bayesian optimization, arxiv preprint (2018) arXiv:1802.02219.
[45] Weiss, S. Method of Preparing O-methyl-isourea Hydrogen Sulfate and O- methyl-isourea Sulfate from cyanamide. US3931316, 1976.
[46] R.A. Sheldon, O.rganic synthesis. Past, present and future, Chem. Ind. (London) 23 (1992) 903–906.
[47] T.T. Qi, G.H. Luo, H.T. Xue, F. Su, J.L. Chen, W.K. Su, K.J. Wu, A. Su, Continuous heterogeneous synthesis of hexafluoroacetone and its machine learning-assisted optimization, J. Flow Chem. 13 (3) (2023) 337–346, https://doi.org/10.1007/s41981-023-00273-1.
[48] R.Z. Liang, X.N. Duan, J.S. Zhang, Z.H. Yuan, B.ayesian based reaction optimization for complex continuous gas-liquid-solid reactions, React. Chem. Eng. 7 (3) (2022) 590–598, https://doi.org/10.1039/d1re00397f.
[49] N. Leonenko, A. Malyarenko, Matérn class tensor-valued random fields and beyond, J. Stat. Phys. 168 (6) (2017) 1276–1301, https://doi.org/10.1007/s10955-017-1847-2.
[50] J. Zhang, N. Sugisawa, K.C. Felton, S. Fuse, A.A. Lapkin, M.ulti-objective Bayesian optimisation using q-noisy expected hypervolume improvement (qNEHVI) for the Schotten–Baumann reaction, React. Chem. Eng. 9 (3) (2024) 706–712, https://doi.org/10.1039/D3RE00502J.
[51] M. Garnelo, J. Schwarz, D. Rosenbaum, F. Viola, D. Rezende, S. Eslami, Y. Teh, Neural processes, arxiv preprint (2018), https://doi.org/10.48550/arXiv.1807.01622.
[52] M. Garnelo, D. Rosenbaum, C. Maddison, T. Ramalho, D. Saxton, M. Shanahan, Y. Teh, D. Rezende, S. Eslami, Conditional neural processes, Int. Conf. Machine Learn. 80 (2018) 1704–1713, https://doi.org/10.48550/arXiv.1807.01613.
[53] H. Kim, A. Mnih, J. Schwarz, M. Garnelo, A. Eslami, D. Rosenbaum, O. Vinyals, Y. Teh, Attentive neural processes, arxiv preprint (2019), https://doi.org/10.48550/arXiv.1901.05761.
[54] C. Finn, P. Abbeel, S. Levine, Model-agnostic meta-learning for fast adaptation of deep networks, Int. Conf. Mach. Learn. 70 (2017) 1126–1135, https://doi.org/10.48550/arXiv.1703.03400.