

A Review of the Applications of Artificial Intelligence in the Process Analysis and Optimization of Chemical Products

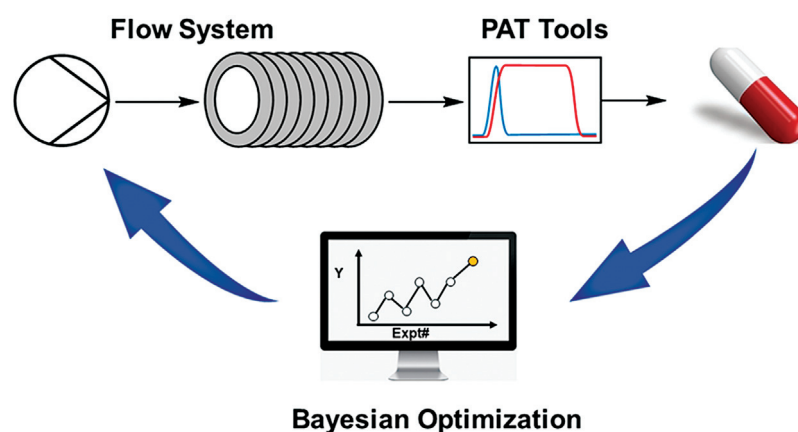
Runqiu Shen¹ Weike Su^{1,2*}

¹Key Laboratory for Green Pharmaceutical Technologies and Related Equipment of Ministry of Education and Key Laboratory of Pharmaceutical Engineering of Zhejiang Province, Collaborative Innovation Center of Yangtze River Delta Region Green Pharmaceuticals, Zhejiang University of Technology, Hangzhou, People's Republic of China

²Zhejiang Yangtze Delta Region Pharmaceutical Technology Research Park, Deqing, People's Republic of China

Address for correspondence Wei-Ke Su, PhD, Collaborative Innovation Center of Yangtze River Delta Region Green Pharmaceuticals, Zhejiang University of Technology, 18 Chaowang Road, Hangzhou 310014, People's Republic of China (e-mail: pharmlab@zjut.edu.cn).

Pharmaceut Fronts 2023;5:e219–e226.



Abstract

Keywords

- ▶ process analysis technology
- ▶ self-optimization
- ▶ artificial intelligence
- ▶ flow chemistry

Continuous flow chemistry is an enabling technology for automated synthesis. Artificial intelligence (AI) is a powerful tool in various areas of automated synthesis in flow chemistry, including process analysis technology and synthesis reaction optimization. The merger of continuous flow chemistry and AI drives chemical production in a more intelligent, automated, and flexible direction. This review discusses the recent application of AI in analyzing and optimizing chemical products produced by continuous flow chemistry with the most innovative equipment and techniques.

received
July 31, 2023
accepted

November 14, 2023
article published online
December 8, 2023

DOI <https://doi.org/10.1055/s-0043-1777425>.
ISSN 2628-5088.

© 2023. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution License, permitting unrestricted use, distribution, and reproduction so long as the original work is properly cited. (<https://creativecommons.org/licenses/by/4.0/>)
Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Introduction

Compared with traditional chemical methods, continuous flow technology has the advantage of allowing a high degree of control over process parameters including temperature, pressure, residence time, ease of scale-up, and automation, and is suitable for multistep synthesis.^{1–3} Thus, it is an ideal technology to control the chemical reactions. Continuous flow chemistry provides an automated-friendly, flexible, innovative, and space-saving reaction platform, and has only recently matured. In recent years, flow chemistry has been involved in an increasing number of experiments. As the amount of one-time access, flow chemistry is especially suitable for dangerous reactions such as diazotization, oxidation, nitrification, etc. As a safe, easily controlled, and green platform, in line with the concept of sustainable development, flow chemistry is receiving increasing attention.

Optimization of synthesis reactions is crucial for both chemical research and discovery. However, optimization, especially in natural chemical production, often involves multiple variables and objectives, making the issue much more complex. To decrease the complexity of the optimization process, chemical automation is preferred and is easily achieved in small-scale continuous flow experiments. Process analytical technology (PAT) is a system for designing, analyzing, and controlling manufacturing processes through the measurement of Critical Process Parameters that affect Critical Quality Attributes (CQAs).⁴ Combining online or inline analytical techniques with flow chemistry helps automate production processes by enabling real-time inspection and process control. For instance, inline nuclear magnetic resonance (NMR) and online infrared (IR) help the system quickly and accurately collect the information desired for production. The information collected is passed on to a computer for processing, which guides this or the next experiment. Rapid and integrated data acquisition through PAT tools allows for data-rich experimentation to be processed by automated optimization algorithms. However, this places a high demand on the device, data acquisition, and processing capabilities of PAT tools. As artificial intelligence (AI) evolves, most of those issues continue to be improved, thereby increasing the efficiency, agility, quality, and flexibility of the current production. PAT tools are the premise and foundation for the self-optimization of AI in flow chemistry. This review summarized the recent AI applications in process analysis and optimization of chemical products produced by continuous flow chemistry.

Process Analytical Technology

Several automated analytical methods have been developed for continuous flow production.^{5,6} For a specific classification of PAT tools and their respective application scenarios and application examples, see Morin et al's work published in 2021.⁷ With PAT tools, experimental data are transferred to AI accurately and quickly for data processing. AI usually starts with data enhancement or dimensionality reduction, using

diverse algorithms such as multi-dimensional scaling to reduce interference and obtain clear and useful data, and this is also known as preprocessing methods (denoising and baseline correction).⁸ For instance, to achieve a machine language conversion of the spectral information, de-spiking, calibration of wavenumber and spectrometer, background correction, and normalization should be performed sequentially.⁹ After the data are standardized, it is processed into machine-readable information through a data analytics model. Therefore, the combination of PAT and intelligent algorithms^{10–12} is powerful in flow chemistry synthesis. This review aims to present the applications of intelligence algorithms in process analysis of continuous flow chemistry synthesis.

In 2020, Kappe's group developed a design of experiments (DOE) model to predict the experimental outcomes and used inline NMR to guide the optimization of a complex nitration reaction in flow.¹³ Low-field NMR suffers from overlapping peaks and quantitative difficulties; however, low-field NMR data can be processed using a multivariate analysis (MVA) with the partial least squares (PLS) regression method. For compounds with spectrally overlapping spectra, the inline NMR tool paired with the MVA method is quite effective despite a very short acquisition time. Previously, data processing methods of PAT were relatively simple and usually used to determine the reaction progress or relative product distribution (i.e., % content). However, more progress is needed in multistep process synthesis combining multiple PAT instruments. Kappe's group applied four different real-time analytical tools (NMR, ultraviolet-visible spectroscopy [UV/vis], IR, and ultra-high-performance liquid chromatography) to a continuous flow multistep synthesis of mesalazine.¹⁴ Mesalazine was obtained from the reactions of nitration, hydrolysis, and hydrogenation, and indirect intricate modeling, deep learning, and PLS regression were created to quantify the desired products, intermediates, and impurities in real-time at different points along the process (→ Fig. 1). These three AI techniques precisely identified product and impurity concentrations, thus controlling the process effectively, and contributes to the optimization and understanding of the reaction.

In 2022, Kappe's group implemented a multi-instrument PAT strategy in a flow line that equips monitors capability for CQAs at several points.¹⁵ They combined the multi-instrument PAT strategy with a DOE strategy to develop process models for all steps of the synthesis of mesalazine, including reaction kinetics, mechanistic models, data-driven based on process data, or depicting residence time distributions.¹⁵ These models contributed to the development of automation concepts to some extent and reveal the full potential of real-time monitoring and model control. In the same year, Kappe's group developed a self-optimizing reactor platform to optimize the reactions efficiently and timely (→ Fig. 2).¹⁶ Input variables are passed to the central supervisory control and data acquisition (SCADA) system in a universal format using Thompson sampling efficient multi-objective (TSEMO) optimization algorithm.^{17,18} The SCADA system transfers the required set points to the peripherals. At the end of the reaction, inline measurements were performed using NMR

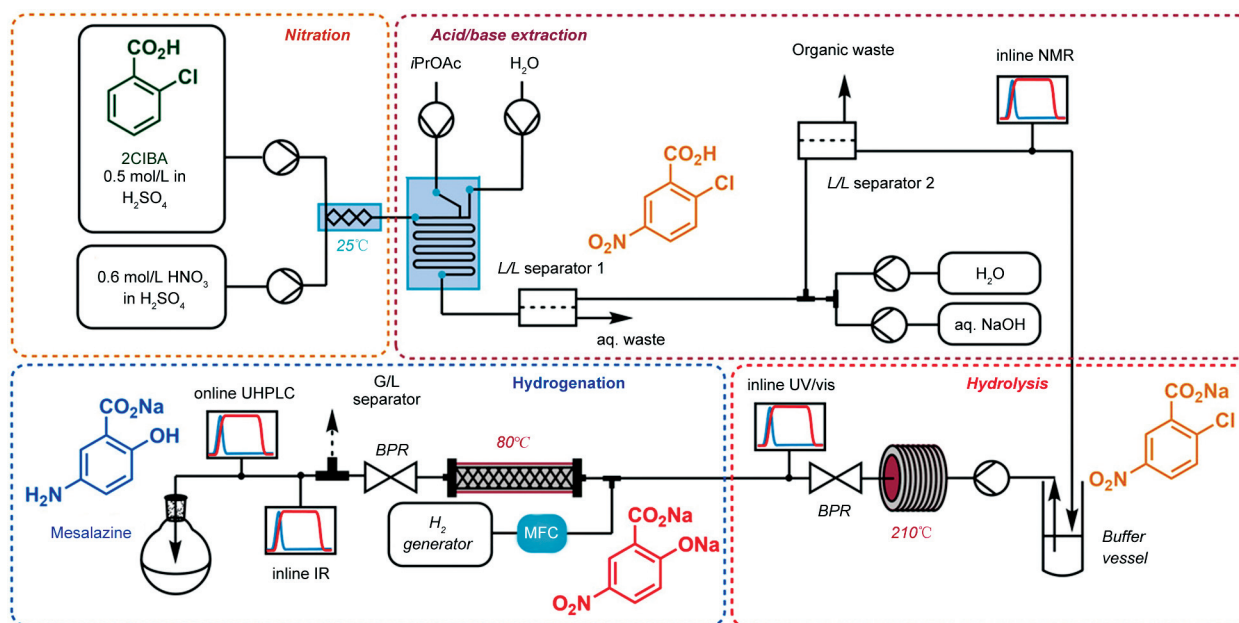


Fig. 1 The multistep, multi-PAT reaction setup toward 5-ASA. Reproduced with permission from a reported study.¹⁴ PAT, process analytical technology; 5-ASA, mesalazine.

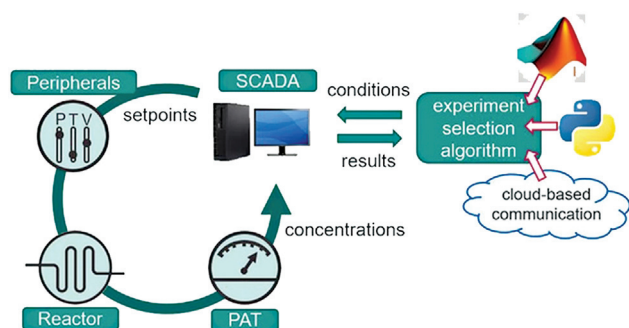


Fig. 2 Overall approach of self-optimizing reactor platform combined with the Bayesian algorithm and inline NMR and FTIR. Reproduced with permission from a reported study.¹⁶ NMR, nuclear magnetic resonance; FTIR, Fourier transform infrared spectroscopy.

and Fourier transform infrared spectroscopy, followed by real-time processing using an indirect hard model and PLS for quantification. The SCADA system further processes the concentration, providing the desired target value, and returning it to the decision algorithm to incorporate the current model and select the next reaction condition.

Furthermore, Kappe's group demonstrated how to use artificial neural networks (ANN) to complete advanced data processing of NMR and UV/Vis spectra to accurately predict the concentration of intermediates in the production of mesalazine.¹⁹ The creation and utilization of simulated training spectra speed up the training process of ANN and compensate for the lack of experimental data. This strategy encourages more use of low-cost and easy-to-get PAT instruments for multistep reaction monitoring. Given the above, PAT provides various analytical data for AI in real-time and plays an important role in flow chemistry.

Intelligent Algorithms for Self-Optimization

Automation has become a reality in life, including chemical synthesis, and it enables scientists to save a great deal of time and effort from time-consuming repetitive tasks. Automatic self-optimization transforms traditional experiment methods designed by scientists into a mode that improves the level of autonomy through intelligent algorithms that enable the system to self-learn, self-decide, self-act, and maximize the experimental efficiency of each design through algorithms that greatly reduce the number of experiments and achieve optimization targets. Automated flow systems quickly search large regions of the experimental space, which makes them suitable for solving optimization problems.^{20,21} Self-optimization combines flow reactors, process analytics, and optimization algorithms to optimize the process. An optimization algorithm receives the responses from the reaction mixture analysis. Based on the results of previous experiments, the algorithm then generates the next set of reaction conditions to be investigated, thus, creating a feedback loop.²²

The development of self-optimization algorithms has a long history, whether it is DOE and Nelder–Mead Simplex (NMSIM) in local optimization algorithms,^{23,24} or Stable Noisy Optimization by Branch and FIT (SNOBFIT)²⁵ in global algorithms, they are all single-objective optimization. However, in complex chemical production, considering various economic and environmental process indicators, single-objective optimization is not suitable for the current situation.²⁶ Therefore, there is a great need to develop an algorithm that can optimize multiple objectives at the same time. Bayesian optimization (BO), as the current mainstream optimization method, is popular to solve the problems mentioned above. BO is mainly composed of a

probabilistic surrogate model, acquisition function, and loss function. A probabilistic surrogate model (e.g., Kriging model and support vector machines) consists of a prior distribution that predicts the likelihood of an unknown objective function and an observation model that describes the data generation mechanism, which can greatly reduce the cost of optimization targets that are high-dimensional, nonlinear, and multi-extreme. Next, the acquisition function predicts the optimal target conditions from the surrogate model and gives guidance for experiments. Then, according to the suggestions given by the acquisition function, the real value obtained by experimentation is compared with the value predicted by the model, and the difference between them is the loss function. So, the smaller the loss function, the better the robustness of the model. Finally, the experimental data are trained again to update the model until a perfect model is obtained. The whole process is called BO. Therefore, in practice, we tend to query the proxy model by minimizing the loss function to. BO can avoid the high computational cost of methods such as DOE (the number of required experiments grows exponentially with the number of factors) and tries to find the optimal solution directly from such variables.

In 2017, Lapkin and coworkers compared the performance of two self-optimizing methods, model-based DOE (MBoE) and Multi-objective Active Learner (MOAL) algorithm, on the C–H activation in flow.²⁷ This study preliminarily explores the application of multi-objective optimization and finds it has great potential in continuous flow. In 2018, Lapkin's group proposed a new multi-objective algorithm, TSEMO, which employed Gaussian processes, spectral sampling, and a genetic algorithm.¹⁷ They studied the nucleophilic aromatic substitution (S_NAr) reaction and N-benzylation reaction, respectively. Both cases optimize four parameters (residence time, substance equivalent, substance concentration, and temperature) at the same time to

maximize the space–time yield (STY) of the reaction while minimizing the E-factor (the proportion of waste).²⁸ The algorithm successfully determines optimal conditions corresponding to a trade-off curve (Pareto frontier) between two chemical reactions in flow. In 2021, Lapkin's group built a self-optimization flow chemistry system with a user-friendly MATLAB user interface.²⁹ Taking the hydroxyaldehyde condensation of benzaldehyde and acetone as an example, the system is initially given a certain amount of experimental data as a training set, and by controlling temperature, flow rate, and equiv. of different components, the Pareto front is achieved. This system is designed to further investigate the optimization behavior of the TSEMO Bayesian optimizer in exploiting and exploring the experimental parameter space. In the same year, Lapkin's group optimized a commercial formulation using robotic experiments driven by a machine learning classification algorithm with the TSEMO algorithm.³⁰ In the experiment, sample preparation, procession, formulation, and analysis were fully controlled and automated by the robot. Then data were then processed by the computer and the variables were optimized by algorithms, forming a highly automated closed loop. After several cycles, the optimal solution is finally obtained. Within 15 working days, the procedure outperforms human intuition in terms of the ability of formulations to achieve the desired result.

Moreover, Lapkin's group developed a framework named Summit to optimize chemical reactions (–Fig. 3).³¹ They presented two benchmarks for reaction optimization and compared the performance of seven strategies on these benchmarks under different combinations of multi-objective transformations. The results show that TSEMO exhibits the best performance in both benchmarks. From here, it can be seen that self-optimization strategies are gradually shifting from the initial exploration to finding better algorithms to

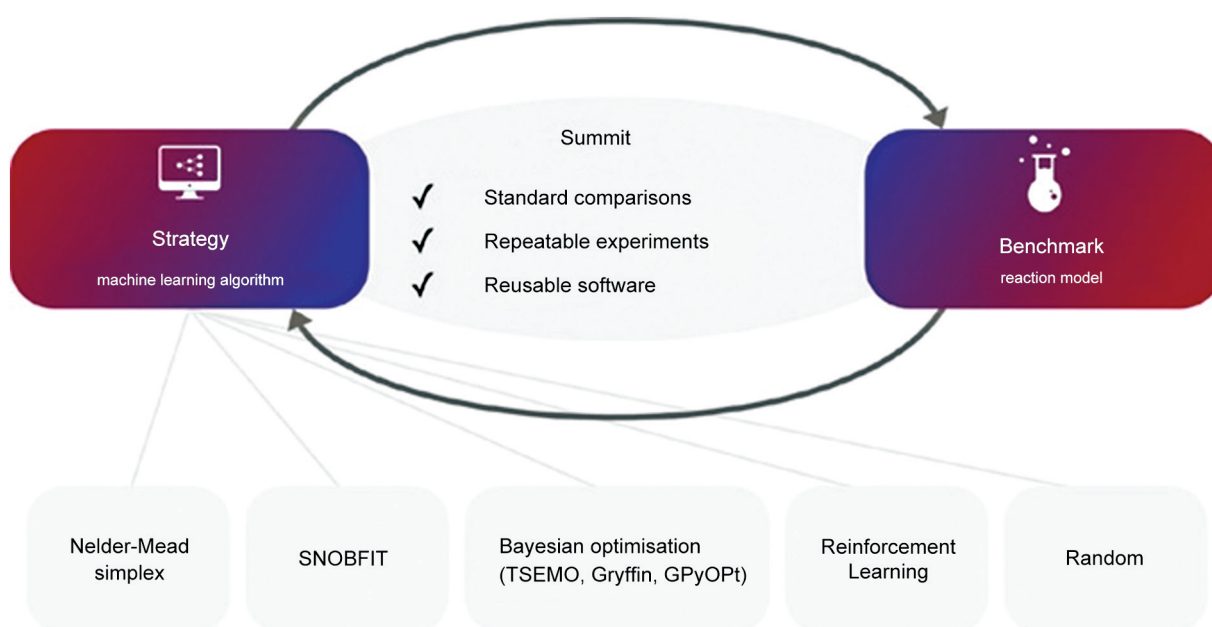


Fig. 3 Overview of the approach used by Summit. Reproduced with permission from a reported study.³¹

achieve low cost and high accuracy. In 2023, Lapkin's group demonstrated a method for adjusting the pH of several multi-buffered polyprotic solutions for formulation chemistry laboratories. Supported by an active optimization strategy based on machine learning, the robotic platform reached the target pH through a wholly automated iterative workflow.³²

In addition to the work of Lapkin's group, many other research groups have recently applied different BO algorithms to different reactions in flow chemistry. In 2020, through the study of the Claisen–Schmidt condensation reaction and the subsequent liquid–liquid separation, Bourne and coworkers used TSEMO for the first time to optimize the multistep reaction and separation processes simultaneously and achieved three objectives.³³ The optimization was initialized with 20 LHC (Latin hypercube sampling, a common sampling method) experiments, followed by 89 experiments designed using the TSEMO algorithm. Out of the 109 experiments, 18 Pareto-optimal solutions were identified and consisted of the Pareto front. This Pareto front allows scientists to visualize information about trade-offs between variables. Sometimes when faced with complex reflection problems, chemical scientists have to distinguish between continuous variables and discrete variables. Discrete variables here can be fixed values under certain environmental conditions, such as material,

reaction solvent, solubility, etc. and often reflected in the algorithm in the form of an integer. In 2021, Bourne' group proposed a mixed variable optimization (MVMOO) algorithm for cases where discrete and discrete variables must be optimized.³⁴ The algorithm offers a cost-effective method for optimizing mixed variable multi-objective optimization problems without the need to reparameterize the discrete variables. In 2022, the group developed an automated polymer synthesis and analysis platform that enables closed-loop multi-objective optimization of RAFT polymerization by combining orthogonal online NMR spectroscopy with inline GPC and TSEMO algorithms.³⁵ On this platform, although there were a few redundant experiments, the automatic polymerization of RAFT, which was not feasible under the original conditions, was successfully realized from the batch tank reaction. Overall, it makes great significance for humans to optimize the conditions for polymerization through no human interaction. Furthermore, the group also developed an automated continuous flow platform and successfully applied it to the Heck cyclization–deprotection reaction sequence (→Fig. 4).³⁶ The coupling of online high-performance liquid chromatography (HPLC) and a BO algorithm with an adaptive expected improvement acquisition function (BOAEI)³⁷ determined a total yield of 81% in just 14 hours. This method represents a readily available technology for accelerating the

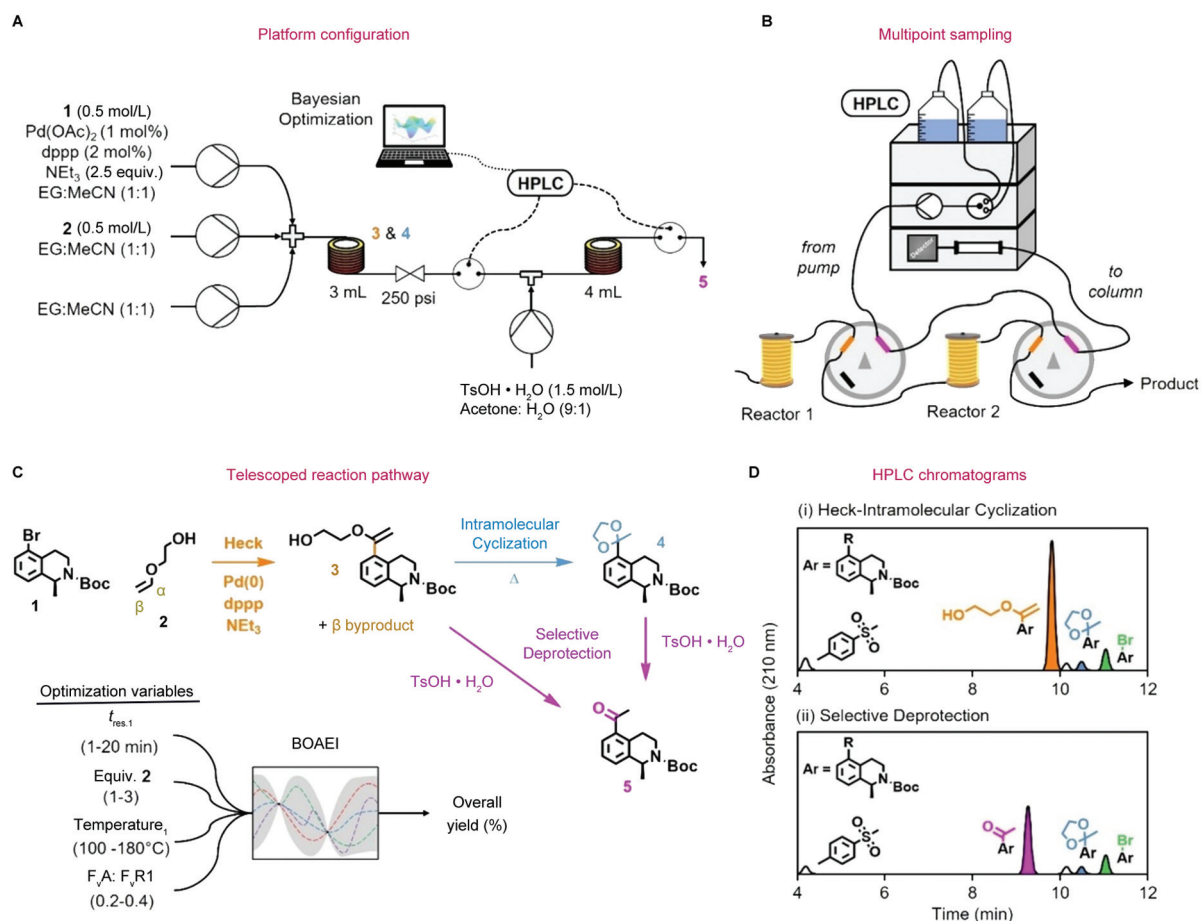


Fig. 4 Self-optimizing reactor platform for multistep synthesis of aryl ketones. Reproduced with permission from a reported study.³⁶

creation of novel chemical syntheses with multiple steps. In 2023, the group applied the MVMOO algorithm to a self-optimizing flow reactor, focusing on S_NAr and palladium-catalyzed Sonogashira reactions.³⁸ The effect of different solvents on the regioselectivity of the S_NAr reactions and the effect of different ligands on the process efficiency of the Sonogashira reactions were determined, and the optimal continuum parameters for both were determined. Specifically, Bourne and coworkers considered solvents as discrete variable and found that common solvents' polarity metrics values as chemical descriptors play a vital role in the S_NAr reaction. In the Sonogashira reactions, the choice of phosphine ligand is regarded as a discrete variable and the trade-offs between common continuous variables and discrete variables were compared. The MVMOO algorithm made it possible to optimize the mixed variable reactions with two objectives effectively. Kappe's group demonstrated an autonomous multi-objective optimization platform for multistep synthesis of edaravone (►Fig. 2).¹⁶ The platform performs multi-objective optimization of the two-step chemical synthesis including seven optimizable variables and three optimization objectives, which makes optimization exponentially more difficult. However, this platform can achieve satisfactory results in relatively few iterations. In terms of the amount of data available and the size of the design space that can be optimized, this is a significant advance in the complexity of autonomous flow reactors.

Kondo et al developed a BO-driven multi-parameter parallel screening to predict reaction conditions and self-optimization process for the flow synthesis of aromatic compounds.³⁹ Through the study of cross-coupling reactions, they successfully and quickly predicted the suitable conditions for the synthesis of 2-amino-2-hydroxy-biaryls (up to 96% yield) and 2,2-dihydroxy biaryls (up to 97% yield). The finally obtained optimized conditions were successfully applied to the synthesis at the gram level.

In recent years, self-optimization strategies have been widely applied in homogeneous reactions or simple heterogeneous reactions. For complex continuous gas-liquid-solid reaction systems, Liang et al developed a continuous reaction optimization platform based on Nelder-Mead simplex method and pure BO algorithm.⁴⁰ In the hydrogenation reaction of nitrobenzene, 3,4-dichloronitrobenzene, and 5-nitro-isoquinoline, BO outperforms only one variable at a time and obtains higher yields and less experimental cost.

Nandiwale et al demonstrated a continuous automated platform for self-optimizing reactions involving solids, which is easy to cause blockage of the reactor channel in flow.⁴¹ The platform consists of a continuous stirred-tank reactor cascade, a slurry feed pump, a photoreactor, and an online HPLC. It is optimized by mixed-integer nonlinear programming,⁴² which is based on the optimal DOE and the sequential response surface method and Dragonfly BO algorithm⁴³ for single and multi-objective optimization, respectively. Experiments showed that Dragonfly BO responds well to both multiple continuous and discrete variables.

Nambiar et al demonstrated an entire process scheme. The Dragonfly BO algorithm first optimized a multistep

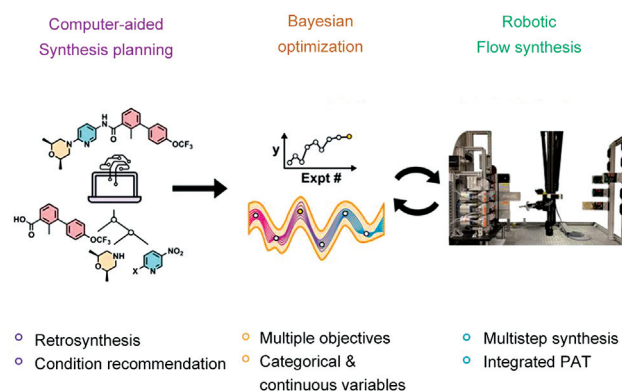


Fig. 5 Overall approach for machine-assisted synthesis planning and process development. Reproduced with permission from a reported study.⁴⁴

synthesis route by AI planning and later ran on an automated robotic flow platform (►Fig. 5).⁴⁴ In contrast to the previously proposed scheme,⁴⁵ the authors introduced a BO algorithm to the multistep synthesis route proposed by the tools of computer-aided synthesis planning, ASKCOS, and PAT to monitor the reaction in real-time on a modular robotic flow synthesis platform. By showing experiments with the full telescoped process for multistep synthesis of sonidegib, they not only demonstrated how the BO works but also identified several areas where human input is still needed such as the measurement of solubility and explanation of the reaction mechanism. Overall, coupling automation, machine learning, and robotics contribute to experiment design, execution, and optimization, which can fuel manual experiments.

Hickman et al used Atinary's transfer learning algorithm SeMOpt, which is a general-purpose, model-agnostic BO framework that uses meta-/few-shot learning to efficiently transfer knowledge from related historical experiments to a novel experimental campaign via a compound acquisition function.⁴⁶ They used the Self-driving Laboratories (SDLabs),⁴⁷ which is also known as Materials Acceleration Platforms (MAPs), to enable knowledge transfer across optimization campaigns by employing neural processes.⁴⁸ After meta-training, SeMOpt performed well in optimizing the simulated cross-coupling reaction and optimizing the Pd-catalyzed Buchwald-Hartwig cross-coupling reaction. Transferred knowledge mainly contains three forms: (1) knowledge transfers from human experts. In other words, scientists manually define the parameters of the selectable area. (2) Knowledge transfers from a proxy campaign. The proxy campaign is chosen to be executed in parallel with a targeted campaign, and the knowledge is dynamically passed from proxy to target when measurements are made in both campaigns. (3) Knowledge transfer from pre-existing campaigns, e.g., knowledge that has been reported before or obtained from open-source databases. This study expresses an important idea that the robot could find usable information from historical correlation data, which can greatly improve the speed of the optimizer.

Dunlap et al optimized flow synthesis of butylpyridinium bromide by using a multi-objective experimental design via a

BO platform (EDBO+).⁴⁹ They used nmrplug, an open-source Python module, to compare semi-automated NMR spectroscopy analysis data and manual processing methods of spectra collected on low-field (60 MHz) and high-field (400 MHz) NMR spectrometers. The optimization objects, reaction yield, and STY were optimized by changing three continuous variables: residence time (T_{res}), temperature (T), and the mole fraction of pyridine (C_{pyr}). The predicted Pareto front has little error compared with the actual experimental data. It shows the power of self-optimization tools for continuous reaction optimization.

As mentioned above, self-optimization in continuous chemistry process development and optimization is an attractive field. However, the establishment of a complete self-optimization platform is still limited by the maturity of the current continuous processes, the advanced nature of the algorithm, and the accuracy of PAT, which requires people to constantly update and improve continuous flow processes, algorithms, and PAT.

In 2022, Hickman and coworkers extended experiment planning algorithms PHOENICS and GRYFFIN,^{50–52} such that they could handle arbitrary known constraints via an intuitive and flexible interface. As GRYFFIN was the more general algorithm, and PHOENICS was included within its capabilities, authors referred only to GRYFFIN. They demonstrated the flexibility and robustness of this algorithm by constraining some continuous and discrete test functions to a certain extent. They also illustrated the algorithm's practical utilities in two simulated chemical research scenarios: the optimization of the synthesis of *o*-xylenyl Buckminsterfullerene adducts under constrained flow conditions and the design of redox-active molecules for flow batteries under synthetic accessibility constraints. In their work, GRYFFIN shows good performance and opens up a broader application of algorithms in chemical reactions.

Conclusion and Outlook

Continuous flow technology has been increasingly emphasized and developed in chemical production and has evolved from a basic organic synthesis to a now important and realistic production method. In this article, we review the successful cases of process analysis and self-optimization applied in AI in continuous flow chemistry. As mentioned earlier, PAT is beneficial for the measurement of various data in continuous chemistry. In addition, AI combined with PAT can model the reaction system to quantitative data of each component in real time and intuitively, effectively improving the chemist's cognitive level of the reaction process. The self-optimization platform is an experimental closed-loop system consisting of PAT and AI, which can automatically detect, analyze, optimize, and quickly converge to find the optimal solution as long as scientists provide original experimental data for the model, thus greatly reducing costs and increasing manufacturing productivity. With the advent of Industry 4.0, AI will be more and more respected. To ensure efficiency, agility, quality, and flexibility in continuous processes, the equipment of PAT and AI algorithms should be constantly

improved. Perhaps in the further, AI could do the tasks entirely independently, meaning that AI has the capacity to acquire basic abilities in a similar manner how human acquires the ability of thinking. Designing experiments, optimizing targets, and obtaining optimal solution sets can be done automatically by AI, ultimately realizing the intelligent and efficient chemical production in the 21st century.

Funding

This work was supported by the Zhejiang Provincial Key R&D Project (Grant No. 2019-ZJ-JS-03).

Conflict of Interest

None declared.

References

- Hartman RL, McMullen JP, Jensen KF. Deciding whether to go with the flow: evaluating the merits of flow reactors for synthesis. *Angew Chem Int Ed Engl* 2011;50(33):7502–7519
- Tsubogo T, Oyamada H, Kobayashi S. Multistep continuous-flow synthesis of (R)- and (S)-rolipram using heterogeneous catalysts. *Nature* 2015;520(7547):329–332
- Sagandira CR, Akwi FM, Sagandira MB, Watts P. Multistep continuous flow synthesis of stavudine. *J Org Chem* 2021;86(20):13934–13942
- Simon LL, Pataki H, Marosi G, et al. Assessment of recent process analytical technology (PAT) trends: a multiauthor review. *Org Process Res Dev* 2015;19(01):3–62
- Fabry DC, Sugiono E, Rueping M. Online monitoring and analysis for autonomous continuous flow self-optimizing reactor systems. *React Chem Eng* 2016;1(02):129–133
- Musio B, Gala E, Ley SV. Real-time spectroscopic analysis enabling quantitative and safe consumption of fluoroform during nucleophilic trifluoromethylation in flow. *ACS Sustain Chem& Eng* 2018;6(01):1489–1495
- Morin MA, Zhang WP, Mallik D, Organ MG. Sampling and analysis in flow: the keys to smarter, more controllable, and sustainable fine-chemical manufacturing. *Angew Chem Int Ed Engl* 2021;60(38):20606–20626
- Houhou R, Bocklitz T. Trends in artificial intelligence, machine learning, and chemometrics applied to chemical data. *Anal Sci Adv* 2021;2(3–4):128–141
- Bocklitz T, Juergen P, Schmitt M. Optical molecular spectroscopy in combination with artificial intelligence for process analytical technology. *Spectroscopy (Springf)* 2020;35(06):28–32
- McGill C, Forsuelo M, Guan Y, Green WH. Predicting infrared spectra with message passing neural networks. *J Chem Inf Model* 2021;61(06):2594–2609
- Lu W, Chen X, Wang L, Li H, Fu YV. Combination of an artificial intelligence approach and laser tweezers raman spectroscopy for microbial identification. *Anal Chem* 2020;92(09):6288–6296
- Lussier F, Thibault V, Charron B, Wallace GQ, Masson J. Deep learning and artificial intelligence methods for raman and surface-enhanced raman scattering. *Trends Anal Chem* 2020;124:115796
- Sagmeister P, Poms J, Williams JD, Kappe CO. Multivariate analysis of inline benchtop NMR data enables rapid optimization of a complex nitration in flow. *React Chem Eng* 2020;5(04):677–684
- Sagmeister P, Lebl R, Castillo I, et al. Advanced real-time process analytics for multistep synthesis in continuous flow. *Angew Chem Int Ed Engl* 2021;60(15):8139–8148
- Sacher S, Castillo I, Rehl J, et al. Automated and continuous synthesis of drug substances. *Chem Eng J* 2022;177:493–501

- 16 Sagmeister P, Ort FF, Jusner CE, et al. Autonomous multi-step and multi-objective optimization facilitated by real-time process analytics. *Adv Sci (Weinh)* 2022;9(10):e2105547
- 17 Bradford E, Schweidtmann AM, Lapkin A. Efficient multiobjective optimization employing gaussian processes, spectral sampling and a genetic algorithm. *J Glob Optim* 2018;71(02):407–438
- 18 Anton SD, Fraunholz D, Lipps C, Poh IF, Zimmermann M, Schotten HD. Two decades of scada exploitation: a brief history. Paper presented at: 2017 IEEE Conference on Application, Information and Network Security (AINS); November 13–14, 2017; Miri, Malaysia
- 19 Sagmeister P, Hierzegger R, Williams JD, Kappe CO, Kowarik S. Artificial neural networks and data fusion enable concentration predictions for inline process analytics. *Digit Discov* 2022;1(04):405–412
- 20 Perera D, Tucker JW, Brahmabhatt S, et al. A platform for automated nanomole-scale reaction screening and micromole-scale synthesis in flow. *Science* 2018;359(6374):429–434
- 21 Richmond CJ, Miras HN, de la Oliva AR, et al. A flow-system array for the discovery and scale up of inorganic clusters. *Nat Chem* 2012;4(12):1037–1043
- 22 Reizman BJ, Jensen KF. Feedback in flow for accelerated reaction development. *Acc Chem Res* 2016;49(09):1786–1796
- 23 Weissman SA, Anderson NG. Design of experiments (doe) and process optimization. A review of recent publications. *Org Process Res Dev* 2015;19(11):1605–1633
- 24 Nelder JA, Mead R. A simplex method for function minimization. *Comput J* 1965;7:308–313
- 25 Huyer W, Neumaier A. Snobfit—stable noisy optimization by branch and fit. *Acm T Math Software* 2008;35(02):1–25
- 26 Jumbam DN, Skilton RA, Parrott AJ, Bourne RA, Poliakoff M. The effect of self-optimisation targets on the methylation of alcohols using dimethyl carbonate in supercritical co₂. *J Flow Chem* 2012;2(01):24–27
- 27 Echtermeyer A, Amar Y, Zakrzewski J, Lapkin A. Self-optimisation and model-based design of experiments for developing a C-H activation flow process. *Beilstein J Org Chem* 2017;13(01):150–163
- 28 Schweidtmann AM, Clayton AD, Holmes N, et al. Machine learning meets continuous flow chemistry: automated optimization towards the pareto front of multiple objectives. *Chem Eng J* 2018;352:277–282
- 29 Jeraal MI, Sung S, Lapkin AA. A machine learning-enabled autonomous flow chemistry platform for process optimization of multiple reaction metrics. *Chem Methods* 2021;1(01):71–77
- 30 Cao L, Russo D, Felton K, et al. Optimization of formulations using robotic experiments driven by machine learning doe. *Cell Rep Phys Sci* 2021;2(01):100295
- 31 Felton KC, Rittig JG, Lapkin AA. Summit: benchmarking machine learning methods for reaction optimisation. *Chem Methods* 2021;1(02):116–122
- 32 Pomberger A, Jose N, Walz D, et al. Automated ph adjustment driven by robotic workflows and active machine learning. *Chem Eng J* 2023;451:139099
- 33 Clayton AD, Schweidtmann AM, Clemens G, et al. Automated self-optimisation of multi-step reaction and separation processes using machine learning. *Chem Eng J* 2020;384:123340
- 34 Manson JA, Chamberlain TW, Bourne RA. Mvmoo: mixed variable multi-objective optimisation. *J Glob Optim* 2021;80(04):865–886
- 35 Knox ST, Parkinson SJ, Wilding CYP, Bourne RA, Warren NJ. Autonomous polymer synthesis delivered by multi-objective closed-loop optimisation. *Polym Chem* 2022;13(11):1576–1585
- 36 Clayton AD, Pyzer-Knapp EO, Purdie M, et al. Bayesian self-optimization for telescoped continuous flow synthesis. *Angew Chem Int Ed Engl* 2023;62(03):e202214511
- 37 Jasrasaria D, Pyzer-Knapp EO. Dynamic control of explore/exploit trade-off in Bayesian optimization. In: Arai K, Kapoor S, Bhatia R, eds. *Intelligent Computing. SAI 2018. Advances in Intelligent Systems and Computing*, vol 858. Cham, Switzerland: Springer; 2019:1–15
- 38 Kershaw OJ, Clayton AD, Manson JA, et al. Machine learning directed multi-objective optimization of mixed variable chemical systems. *Chem Eng J* 2023;451:138443
- 39 Kondo M, Wathsala HDP, Salem MSH, et al. Bayesian optimization-driven parallel-screening of multiple parameters for the flow synthesis of biaryl compounds. *Commun Chem* 2022;5(01):148
- 40 Liang R, Duan X, Zhang J, Yuan Z. Bayesian based reaction optimization for complex continuous gas–liquid–solid reactions. *React Chem Eng* 2022;7(03):590–598
- 41 Nandiwale KY, Hart T, Zahrt AF, et al. Continuous stirred-tank reactor cascade platform for self-optimization of reactions involving solids. *React Chem Eng* 2022;7(06):1315–1327
- 42 Baumgartner LM, Coley CW, Reizman BJ, Gao KW, Jensen KF. Optimum catalyst selection over continuous and discrete process variables with a single droplet microfluidic reaction platform. *React Chem Eng* 2018;3(03):301–311
- 43 Kandasamy K, Vysyaraju KR, Neiswanger W, et al. Tuning hyperparameters without grad students: scalable and robust Bayesian optimisation with dragonfly. *J Mach Learn Res* 2020;21(01):3098–3124
- 44 Nambiar AMK, Breen CP, Hart T, Kulesza T, Jamison TF, Jensen KF. Bayesian optimization of computer-proposed multistep synthetic routes on an automated robotic flow platform. *ACS Cent Sci* 2022;8(06):825–836
- 45 Coley CW, Thomas DA III, Lummiss JAM, et al. A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science* 2019;365(6453):6453
- 46 Hickman RJ, Ruža J, Tribukait H, Roch LM, García-Durán A. Equipping data-driven experiment planning for Self-driving Laboratories with semantic memory: case studies of transfer learning in chemical reaction optimization. *React Chem Eng* 2023;8:2284–2296
- 47 Coley CW, Eyke NS, Jensen KF. Autonomous discovery in the chemical sciences Part I: progress. *Angew Chem Int Ed Engl* 2020;59(51):22858–22893
- 48 Kim H, Mnih A, Schwarz J, et al. Attentive neural processes. Paper presented at: ICLR 2019 New Orleans, LA, United States, May 6–9, 2019
- 49 Dunlap JH, Ethier JG, Putnam-Neeb AA, et al. Continuous flow synthesis of pyridinium salts accelerated by multi-objective Bayesian optimization with active learning. *Chem Sci (Camb)* 2023;14(30):8061–8069
- 50 Hickman RJ, Aldeghi M, Häse F, Aspuru-Guzik A. Bayesian optimization with known experimental and design constraints for chemistry applications. *Digit Discov* 2022;1:732–744
- 51 Häse F, Roch LM, Kreisbeck C, Aspuru-Guzik A. Phoenix: a Bayesian optimizer for chemistry. *ACS Cent Sci* 2018;4(09):1134–1145
- 52 Aldeghi M, Häse F, Hickman RJ, Tamblyn I, Aspuru-Guzik A. Golem: an algorithm for robust experiment and process optimization. *Chem Sci (Camb)* 2021;12(44):14792–14807