A Hybridised Optimisation of an Automated Photochemical Continuous Flow Reactor

Abstract: A new hybridized algorithm that combines process optimisation with response surface mapping was developed and applied in an automated continuous flow reaction. Moreover, a photochemical cascade CSTR was developed and characterised by chemical actinometry, showing photon flux density of ten times greater than previously reported in batch. The success of the algorithm was then evaluated in the aerobic oxidation of sp² C–H bonds using benzophenone as photosensitizer in the newly developed photo reactor.

Keywords: Algorithm · Automation · Continuous flow · Photochemistry · Self-optimisation

1. Introduction

Modern society’s expectations demand a new level of chemical development, which aligns with sustainability principles and the molecular complexity required across a range of end-use applications. The development of new technologies and methods have a critical role in achieving these needs, improving research output and enabling discovery and evolution of new and strategically important processes. Continuous flow methods have contributed to numerous improvements over batch alternatives, which include: precise reaction times, enhanced heat and mass transfer, access to an extended operating window, and the use of heterogeneous reagents and catalysts in packed-bed reactors.[1–7]

Photons are often considered as green and traceless reagents to instigate chemical reactions, as they can replace chemical activating agents, thereby reducing by-product formation and reducing the downstream processing required. For these reasons, there has been a recent uptake in the use of micoreactors for photochemistry, with the narrow channels providing a high surface area to volume ratio. This results in a uniform distribution of light across the reactor, increasing selectivity and reducing reaction times.[18] However, limitations still exist regarding the scale-up of these systems. Although numbering-up has been previously demonstrated,[19] there remains operational challenges in the context of highly regulated manufacturing. Therefore, continuous stirred tank reactors (CSTRs) remain one of the favoured choices of reactor for high productivity applications in the pharmaceutical industry.[10] One example application is the development of a 100 mL laser-driven CSTR for a kilogram-scale photocatalysed C–N coupling reaction, where a high intensity laser (25 W) was utilized to increase the total amount of light absorbed, and thereby the productivity.[11] The recent developments in LED technology have reduced the entry barrier to using photochemical reactions and ignited a renewed interest in photochemical methods.[12,13] This has led to a new generation of photochemical reactors and methods.[8,12–16]

Continuous flow processes lend themselves to automation with the availability of in-line and on-line analytical techniques.
To maximise efficiency these automated systems can be coupled with optimisation algorithms to close the feedback loop. This allows for the opportunity to free the human resource associated with experimentation and significantly improve the productivity and efficiency of laboratory and process development. With the drive from pharmaceutical industry to adopt a quality by design (QbD) approach there is a requirement to understand the chemical reaction domain. Current techniques have focused on adopting a statistical or mechanistic approach to this problem through iterative screening and optimisation or postulating a kinetic model. As these approaches can often require extensive data collection, there is a growing need for the development of intelligent algorithms. In this work, we present a hybrid algorithm which locates the optimum and efficiently maps the local response surface terrain. Notably, this is achieved in a significantly reduced number of experiments compared to successive DoEs which have been extensively used by industry for many years.

2. Results and Discussion

2.1 Miniature Photochemical Continuous Stirred Tank Reactor (CSTR) Cascade

We have previously reported a miniature CSTR cascade (known as Freactor\textsuperscript{[23,24]} for the laboratory-scale optimisation of multiphasic chemical reactions.\textsuperscript{[21]} Active mixing provided by the CSTRs decouples flow rate and mixing performance, making them suitable for mass transfer limited reactions with longer residence times (> 10 min). This also enables a low liquid hold-up to be maintained, which is desirable for minimising material consumption during process development (e.g. self-optimisation). Taking this into consideration, we designed a photochemical miniaturised CSTR cascade for laboratory-scale process development (Fig. 1, where \( n \) is the number of CSTRs). Furthermore, the agitation provided by a CSTR prevents the settling of particulates, thus reducing the risk of reactor fouling.\textsuperscript{[26]} This is desirable for maintaining a constant reactor performance, as precipitation on the reactor walls reduces the amount of light absorbed. Due to the higher cost and safety concerns regarding the use of lasers, we opted for commercially available LEDs (365 nm, 3 W nominal power per CSTR) to provide a more widely accessible system. Mixing within the CSTR transports fluid vertically and thus into and out of the region close to the window where the flux density is greatest. This active mixing overcomes the issue of diminishing light intensity as a function of distance travelled through the reaction medium and overcomes the diffusion controlled transport of reactants and products into and out of the photochemically active region.

![Fig. 1. Photograph of the photochemical miniature CSTR cascade (\( n = 4 \)) used in the setup for the optimisation of tetralin to tetralone (Fig. 6).](image)

2.2 Chemical Actinometry

The photon flux density defines the amount of light absorbed per unit of volume per unit of time. The higher the photon flux density of a reactor, the faster a photochemical process will be complete. To calculate this, the incident light intensity (\( I_0 \)) must be determined. This was achieved via chemical actinometry, which uses a photoinduced reaction of a compound to measure the light intensity at a given wavelength. The photochemical isomerisation of \( o \)-nitrobenzaldehyde (NBA) to \( o \)-nitrosobenzoic acid at 365 nm was investigated using an automated continuous flow reactor (Fig. 2). Notably, the quantum yield (\( \phi \)) of the chemical actinometer must be known. In this case, the quantum yield of NBA was known to be 0.5 at 365 nm.\textsuperscript{[27]}

![Fig. 2. Schematic of the setup for automated chemical actinometry (\( n = 2 \)). \( o \)-Nitrobenzaldehyde was used as a chemical actinometer with known quantum yield (\( \phi = 0.5 \)). The photochemical isomerisation of \( o \)-nitrobenzaldehyde to \( o \)-nitrosobenzoic acid was monitored via HPLC analysis. SL = sample loop.](image)

The relationship between the rate of conversion of the actinometric compound and the Beer-Lambert law can be used to derive Eqn. (1) for the calculation of the incident light intensity. The experiments are conducted with high concentrations ([NBA] = 0.1 M) of the actinometric compound and long path lengths (\( l = 1 \) cm). This makes the right hand term of Eqn. (1) close to unity, and enables the negative differential term to be replaced by a zero-order rate constant. Therefore, Eqn. (1) can be simplified to Eqn. (2).\textsuperscript{[22,23]}

\[
I_0 = \frac{d[\text{NBA}]}{dt} \left( \frac{k_0}{\phi} \right) \left( \frac{1}{1 - 10^{-\phi[NBA]}} \right) \quad (1)
\]

\[
I_0 = \frac{k_0}{\phi} \quad (2)
\]

The zero-order rate constant is equal to the slope of the time profile for the photochemical conversion of NBA (Fig. 3). However, significant curvature was observed as a result of light absorption by the \( o \)-nitrosobenzoic acid product. Therefore, a second-order polynomial was fitted to the data and the initial slope of the curve determined by taking the derivative and evaluating at time = 0 min. The resultant zero-order rate constant (\( k_0 = 1.67 \mu g \mu L^{-1} \text{min}^{-1} \)) was used to calculate the incident light intensity (\( I_0 = 3.67 \times 10^3 \text{ einstein L}^{-1} \text{s}^{-1} \)) according to Eqn. (2), which is equivalent to the photon flux density (\( q/V = 0.37 \text{ einstein m}^{-3} \text{s}^{-1} \)) of the reactor under operation at full absorption. The photon flux density
The dynamic nature of the system was a requirement, with the consumption. The authors were able to generate a polynomial model to minimise the efficiency of data collection in minimising material response surface models utilising dynamic experiments to maximising the yield. The final stage of the algorithm performs a central composite design of the system inputs to ensure system inputs are kept within a desired range.

2.3 Hybrid Algorithm

The coupling of algorithms can lead to enhanced capability and performance simply not accessible by using a single algorithm. In regions where the response surface is expected to be smooth, local search heuristics can outperform their global counterparts in the search for a defined target area. Given this, the implementation of hybrid algorithms could provide an interesting area of development for chemical reaction optimisation.

Determining the optimum for a process is essential to enable efficient process operation. However, it is important to understand the response surface around the optimum, given the dynamic nature of chemical processes. Full knowledge of the response surface around an optimum can inform the user of how changes in process inputs, either deliberate operator inputted changes or changes due to plant error such as pump flowrate or reactor temperature drift, affect the response variable for the system. With single point optimisation, should no further response work be performed, the user lacks full understanding of whether the optimum determined is near a performance ‘cliff edge’ or whether there is a larger unexplored region of process stability. Response surface work around the optimum allows the user to ascertain this knowledge which can further inform the continued design process. This is a vital aspect when adopting a QbD approach, where the limits of operation around the processing conditions must be known to provide a rigorous plant-scale process.

Previous work has focused predominantly on the determination of an optimum set of conditions in an efficient and robust manner.[31] Further understanding of the area around the optimum has not generally been considered as a requirement for optimisation algorithms. In systems where a full mechanistic understanding is still yet to be obtained, it is vitally important to have this information, to ensure system inputs are kept within a desired range.

Recent work by Wyvratt et al. has looked at development of response surface models utilising dynamic experiments to maximise the efficiency of data collection in minimising material consumption. The authors were able to generate a polynomial model describing the response surface for a Knoevenagel condensation reaction. Post processing of the data to correct for the dynamic nature of the system was a requirement, with the primary goal of the methodology to understand a predefined reaction space.[12]

Currently, there is little in the area of optimisation techniques coupled with response surface mapping when applied to self-optimising chemical systems. As such, a hybrid algorithm considering these factors is proposed. Fig. 4 illustrates the conceptual goal of the algorithm, initially determining the process optimum (red point) and then iteratively understanding a user-defined region around the optimum. The hybrid algorithm we have developed is separated into three stages: optimisation, screening and fine-tuning the response surface.

SNOBFIT[33] was selected as the initial optimisation algorithm for the hybrid approach. The algorithm is robust in the presence of noise and has inbuilt exploration of the experimental domain to provide confidence in the determined solution. The exploratory nature of the SNOBFIT algorithm has previously been utilised to fit an empirical polynomial model, with the empirical models showing excellent fit to the data.[19]

Following the initial global optimisation, the algorithm moves onto the screening stage, where a surrogate Gaussian process (GP) model is constructed. A GP model was selected ahead of a polynomial model due to the increased flexibility afforded and the improved capabilities of the GP model to handle noisy data. Fitting of the GP model was preformed using MATLAB’s inbuilt Gaussian process regression function ‘fitrgp’. This fits a Gaussian process model to the available data through minimising a cost function; the cost function balances how well the model fits to the data with its predictive power in relation to future data. Employing a Plackett-Burman screening design, a screening process is undertaken where the upper and lower bounds of the design are optimised to explore the predefined target region. The target provided by the user directs the algorithm to mapping the space around the optimum up to a predefined region. Fig. 4 provides an example of this, with the algorithm aiming to map the region spanning 10% from the optimum yield. As this region is around the previously determined global optimum, a local search method (MATLAB’s ‘fmincon’ function) was utilised to optimise the design.

The optimisation of the design leads to a set of candidate experiments where the chemical system is to be evaluated. This process continues until the screening portion of the algorithm achieves the target defined by the user at initialisation. The final stage of the algorithm performs a central composite face centred (CCF) design utilising the upper and lower bounds determined from the screening portion of the algorithm. This stage of the algorithm further refines the response surface model around the previously ascertained global optimum. The stages undertaken by the algorithm are highlighted in Fig. 5.
2.4 Automated Reaction Optimisation of the Continuous Aerobic Oxidation Using a Hybrid Algorithm

The aforementioned hybrid algorithm was combined with an automated flow reaction system and applied to the aerobic oxidation of tetralin (Scheme 1). The reaction was modified from the original paper[10] to be performed with air at atmospheric pressure. Furthermore, the tetra-n-butylammonium decatungstate (TBADT) photocatalyst was replaced with benzophenone as a more available and atom economical alternative.

The reaction was optimised to maximise tetralone yield, utilising the reactor setup highlighted in Fig. 6. As previously discussed the initial global optimisation was performed utilising the SNOBFIT algorithm, with a subsequent surrogate GP model built from its results. A target tetralone yield reduction of 10% from the optimum value was selected prior to running the algorithm. Product composition for the above scheme was determined by online HPLC, from which the data was used as feedback for the hybrid algorithm. The design space for the hybrid algorithm considered two variables: residence time and oxygen equivalents (Table 1).

Table 1. Global optimisation limits for the tetralin aerobic oxidation. Reservoir solutions: tetralin = 0.04 M in ethyl acetate, benzophenone = 0.5 equiv.

<table>
<thead>
<tr>
<th>Global Limits</th>
<th>Residence Time (min)</th>
<th>Oxygen Equivalents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Upper</td>
<td>30</td>
<td>5</td>
</tr>
</tbody>
</table>

The automated reactor setup was started in the morning and terminated the following evening, running overnight. The algorithm automatically terminated when the final CCF had been completed with a total of 61 experiments performed divided as follows: 38 SNOBFIT experiments, 12 screening experiments and 11 CCF experiments. A graphical summary of the results for the first two stages of the algorithm is provided in Fig. 7, with the final screening design summarised as a contour plot in Fig. 8.

A final response surface model for the entire process (colour mapped surface) can be generated utilising the experimental data (Fig. 9). The plot includes the 95% confidence interval (CI) for the entire search domain (contained within the pink surfaces). A global optimum of 65% tetralone yield was determined with a residence time of 18.3 minutes and 4.69 oxygen equivalents. This compares favourably against previously reported conditions:[6] (i) reduced residence time (18.3 cf. 45 minutes) provides a more productive process; (ii) air is a safer source of oxygen compared to pure oxygen; (iii) benzophenone is a more accessible and atom economical photosensitiser compared to TBADT, even when used at 0.5 equivalents. These benefits can be in part associated with the rapid mixing provided by the CSTRs, which helps to overcome mass transfer limitations associated with gas-liquid segmented flow observed in mesoscale tubular reactors. As a covariance function with automatic relevance determination (ARD) was used, examination of the relevant importance of a variable can be carried out a posteriori (Table 2).

Table 2. Hyperparameters of the final GP model considering all data. Lower value indicates greater importance.

<table>
<thead>
<tr>
<th>Hyperparameter</th>
<th>Relative Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residence Time</td>
<td>13.7184</td>
</tr>
<tr>
<td>Oxygen Equivalents</td>
<td>4.0281</td>
</tr>
</tbody>
</table>

The strong dependence on oxygen equivalents shown by the hyperparameters of the model is clearly indicated in the data and the contour plot (Fig. 7 and Fig. 8). This, coupled with the full model (Fig. 9) provides both qualitative and quantitative results that can be used to inform the design of a chemical process. The accompaniment of a 95% CI can additionally guide this process by offering the expected range of values, imperative in chemical process design. In data dense regions the 95% CI can be seen to be narrower indicating a greater model confidence in the mean value for these regions. Both the traces and the surface model indicate the possibility of over oxidation of tetralone, leading to a...
reduction in yield at longer residence times. Over oxidation in this reaction had previously been observed in the original paper and was confirmed through offline analysis (see ESI for details). We were unable to prepare analytically pure over oxidation products to calibrate the online HPLC due to their decomposition upon separation. Nevertheless, the mass balance for the reaction was maintained above 90% suggesting only small quantities of side-product formation.

The global search nature of SNOBFIT is clearly indicated by the spread of experiments, with the algorithm attempting to build a complete picture of the design space. The screening portion of the algorithm then systematically searches for regions near the target value based on an internal local optimisation of the surrogate GP model. Analysing Fig. 7, it is clear that although the screening portion is able to ascertain the region where a target reduction has occurred, it fails to effectively explore certain areas (longer residence times) where the yield is likely to lie within the target reduction.

3. Conclusions

In summary we have developed a photochemical cascade CSTR reactor and fully characterised it through chemical actinometry, showing it to have a photon flux density ten times greater than previously reported in batch photochemical reactors. The cascade reactor was subsequently demonstrated in the aerobic oxidation of tetralin, with the experimental method automated and optimised using a hybrid algorithm that combined process optimisation with response surface mapping. An optimum tetralone yield of 65% was achieved with an 18.3 minute residence time and 4.69 oxygen equivalents. To the best of our knowledge this is the first time a hybrid algorithm combining two synergistic aims has been applied to the automated optimisation of a chemical system. The development of the hybrid approach described here opens exciting opportunities to increase robustness of chemical processes and operate with an improved understanding of conditions surrounding the optimum.

**Supplementary Information**

Supplementary information is available on https://www.ingentaconnect.com/content/scs/chimia

**Acknowledgements**

JM (EP/R513258/1), AC (EP/N509681/1) and RL (EP/R032807/1) thank the EPSRC and University of Leeds for funding. AC also thanks AstraZeneca for CASE student funding.