Optimisation of Discrete and Continuous Parameters for the Manufacture of Pharmaceuticals

Oliver J. Kershaw

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Chemical and Process Engineering

September 2023

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The work in Chapter 2 and 3 has appeared in: **O. J. Kershaw**, A. D. Clayton, J. A. Manson, A. Barthelme, J. Pavey, P. Peach, J. Mustakis, R. M. Howards, T. W. Chamberlain, N. J. Warren and R. A. Bourne, *Chem. Eng. J.* **2023**, 451, 138443. OJK was responsible for carrying out all experimentation for both case studies, system optimisation and data analysis including data visualisation and simulation. The contributions from other authors were writing the optimisation algorithm, coding, system support and preparation of the manuscript (JAM, ADC) and project supervision (AB, JP, PP, JM, RMH, TWC, NJW, RAB).

“Great things are not done by impulse, but by a series of small things brought together” – Vincent Van Gough

Acknowledgements

Firstly, I would like to thank my supervisors Richard Bourne and Nicholas Warren for supporting and advising me throughout my project, especially in the toughest of times (the fume hood incident). Although the constraints of isolations hindered in-person meetings, the guidance each of you provided has been extremely helpful in enabling me to excel in my professional endeavours. I am exceptionally grateful to have had your support and supervision throughout this project, and for the pathways I chose to follow.

I also wish to thank everyone that I have had the pleasure to meet and collaborate with throughout the duration of this project. To the team at Pfizer (Roger Howard and Phillip Peach) who have helped with their keen insight and support into the project throughout all the meetings. To Jamie Manson, your contribution and support could not be appreciated more, you have helped me through the hardships, especially when coding-related, where your algorithm has allowed this project to take place! I wish to specially thank Adam Clayton-in the later stages of this project, you have become a great mentor for guidance and direction. I would also like to extend my gratitude to Tom Shaw, Joe Houghton and Holly Clarke who have been a source of sanity and support outside of the project as well as introducing me to climbing. To Tom Dixon and Sarah Boyall, I would like to extend my appreciation for their efforts in transforming the office into a social and enjoyable workspace. Additionally, I would like to thank all the current and former members of the iPRD lab, who have made my time at Leeds a pleasant experience: Ricardo, Ilias, Connor, Calum, Luke, Brendan, Mary, Andrew, Zara, Joe and Dan.

I would like to extend my heartfelt appreciation to my entire family, Charlotte, Dan and Abbie Kershaw, with special gratitude for my parents, Anne and David Kershaw. Your unwavering support and affection throughout the entirety of this project has been insurmountable. Your constant encouragement has been a driving force behind my motivation during these studies.

Finally, to my loving partner Bethan, without your continual motivation I would not have realised my current opportunities and determination. You have been amazing throughout, even providing me with a beautiful dog.

Abstract

Continuous flow processing has revolutionised the field of chemistry by enabling enhanced heat and mass transfer, safer handling of hazardous reagents, end-to-end processing of telescoped reactions and access to a broader range of reaction conditions compared to traditional batch methodologies. The desire to reduce the labour and material demands in research and development (R&D) processes has emphasised the need for automation in chemical synthesis. Flow platforms have played a pivotal role in enabling the automation of chemical systems, offering enhanced control over reaction parameters. Consequently, this has led to their substantial integration into the pharmaceutical industry.

The implementation of algorithms in feedback loops on automated flow platforms has expanded the potential of these systems, facilitating efficient exploration and optimisation of chemical processes. This synergy has resulted in the development of proficient self-optimisation systems that adeptly navigate experimental domains, expediting the discovery of optimal conditions and enhancing process understanding. Expanding the scope to encompass all factors in experimental self-optimisation approaches will promote more extensive utilisation, contributing to the advancement of sustainable practices in early-stage reaction development within the pharmaceutical industry.

The work in this thesis aims to unlock the potential of self-optimisation flow platforms, extending the capabilities into previously unexploited areas within this field. This involves introducing discrete variables into automated self-optimisation processes, applying them in the synthesis of a TRPV1 receptor antagonist API and extending the approach to incorporate telescoped flow reactions with consideration of multiple objectives to highlight the effectiveness of end-to-end optimisations.

Table of Contents

Acknowledgements ..... **iv**

Abstract ..... **v**

Table of Contents ..... **vi**

List of Figures ..... **viii**

List of Schemes ..... **xiv**

List of Tables ..... **xvi**

List of Abbreviations ..... **xviii**

Chapter 1. Introduction **22**

1.1 Fundamentals of Continuous Flow 23

1.2 Advantages of Flow Chemistry 26

1.2.1 Containment of Hazardous Chemicals 26

1.2.2 Mixing and Mass Transfer 28

1.2.3 Unlocked Temperatures and Pressures 30

1.2.4 Telescoped Synthesis 31

1.3 Automated Self-Optimisation 34

1.3.1 Local Optimisation 36

1.3.2 Global Optimisation 47

1.4 Mixed Variable Systems 55

1.5 Discussion 59

1.6 Project Aims 60

Chapter 2. Synergising Simulation and Experimental Studies for Mixed Variable Optimisation **62**

2.1 Introduction 63

2.2 Mixed Variable Multi-Objective Optimisation Algorithm 68

2.3 Nucleophilic Aromatic Substitution 71

2.4 Simulated Optimisations 72

2.4.1 S_NAr Simulation 72

2.5 Experimentally Optimised Results 81

2.6 Conclusion 91

Chapter 3. Exploring API Synthesis: A Mixed Variable Optimisation Approach of Catalytic Systems **93**

3.1 Introduction 94

3.2 Sonogashira Cross-Coupling Reaction 98

3.3 Automated Flow Platform and Optimisation Results 101

3.4 Conclusion 114

Chapter 4. Enhancing Telescoped Chemical Reactions through Mixed Variable Optimisation 117

4.1 Introduction 118

4.2 Heck Reaction 122

4.3 Telescoped Optimisation 126

4.4 Conclusion 140

Chapter 5. Conclusions and Future Work 142

Chapter 6. Experimental 148

6.1 Discrete Variable Automated Flow Platform 149

6.2 Offline Analytical Equipment 151

6.3 Chapter 2 Procedures 151

6.3.1 Chemicals 151

6.3.2 Synthesis of *ortho*-2.3, *para*-2.4 and bis-2.5 152

6.3.3 Simulated Optimisation 153

6.3.4 Experimental Set-Up 164

6.3.5 Self-Optimisation Results 167

6.4 Chapter 3 Procedures 172

6.4.1 Chemicals 172

6.4.2 Synthesis of 2-(3,3-dimethylbut-1-yn-1-yl)-4-(trifluoromethyl)benzonitrile 3.6 173

6.4.3 Experimental Set-Up 174

6.4.4 Self-Optimisation Results 178

6.5 Chapter 4 Procedures 182

6.5.1 Chemicals 182

6.5.2 Synthesis of 2-methyl-2-phenyl-1,3-dioxolane 4.8 183

6.5.3 Experimental Set-Up 183

References 192

List of Figures

Figure 1. Comparison of reaction coordinates between flow technologies and conventional batch methodologies 24

Figure 2. Photo-flow platform used in the optimisation utilising a four miniature CSTR cascade for the aerobic oxidation of tetralone from tetralin with benzophenone as a photo synthesiser. 26

Figure 3. (a) Example of Laminar flow in tubing where there is no disruption between parallel fluid layers. (b) Turbulent flow where there is a presence of rapid changes in velocity. 29

Figure 4. The telescoped flow platform employed in the synthesis of Tamoxifen is equipped with various temperature zones for reactions, in addition to pre-mixing and quenching coils.[23] 34

Figure 5. Illustrative example of a self-optimisation flow platform with an automated feed-back loop for system control and direction of conditions. 36

Figure 6. Example two variable OVAT optimisation leading to a false optimum to be observed. ★ represents the best results observed from the process. 37

Figure 7. Example 2-level (*-1, +1*) 3-factor FFD (*X1, X2, X3*) with a singular centre point. The constructed matrix for the experiments is highlighted on the right. Blue points represent experiments required from this matrix, with the centre point highlighted in red, originating from the data entry for experiment 9. 40

Figure 8. Possible geometric transformations performed by the Nelder-Mead Simplex algorithm for a 2D problem: multiple contraction (MC), inside contraction (X_{IC}), outside contraction (X_{OC}), reflection (X_R), expansion (X_e). 42

Figure 9. A visual example of a 2-variable simplex optimisation converging on a local optimum where a global optimum of a better solution is present. 47

Figure 10. Flow diagram for a SNOBFIT optimisation. Where nreq = minimum number of required function evaluations to be performed, set by the user, neval = number of actual function evaluations performed. 49

Figure 11. An example of a conflicting objective optimisation, between productivity/STY and Yield, where both metrics are being minimised. Red points illustrate the Pareto front between the objectives, with blue points being dominant solutions. 52

Figure 12. Hypervolume plot for a 2D objective space. The hypervolume is calculated between the reference point, *r*, and the current non-dominated solutions, s_i. 53

Figure 13. Flow chart for the TSEMO algorithm workflow within an optimisation process, utilising random sampling of GP surrogate models. 55

Figure 14. Photo of a Chemspeed HTE batch reactor, where aliquots of mixtures are drawn up using syringe pumps and injected into the batch vessels. 56

Figure 15. (a) Schematic of the autosampler-flow platform used to screen libraries of discrete variables for the Suzuki-Miyaura cross-coupling reaction. (b) Preparation and injection of discrete variable slugs into the flow stream, where mixing and diffusion it analysed by UV detection. 57

Figure 16. Comparison between traditional HTS methodologies vs. this developed MVMOO self-optimisation approach for a full reaction optimisation pathway. 66

Figure 17. Flow chart for the iterative process of the MVMOO algorithm. 70

Figure 18. SNOBFIT simulation for the S_NAr reaction with pyrrolidine with the *ortho* product selected as the objective. The colour bar for this graph represents the yield for the *ortho* product, with a range of 0-100%. Where the ★ represents the optimum *ortho* yield. 77

Figure 19. SNOBFIT simulation for the S_NAr reaction with pyrrolidine with the *para* product selected as the objective. The colour bar for this graph represents the yield for the *para* product, with a range of 0-4%. Where the ★ represents the optimum *para* yield. 77

Figure 20. SNOBFIT simulation for the S_NAr reaction with morpholine where the *ortho* (2.3) yield is selected as the objective. Where the ★ represents the optimum *ortho* yield. 78

Figure 21. SNOBFIT simulation for the S_NAr reaction with morpholine where the *para* (2.4) yield is selected as the objective. Where the ★ represents the optimum *para* yield. 79

Figure 22. Illustrates the results of a multi-objective optimisation approach using a genetic algorithm to optimise the S_NAr reaction. The objectives of this optimisation were to maximise the yields of the *ortho* (2.3) and *para* (2.4) products from the reaction. 80

Figure 23. Automated flow platform utilised within the solvent optimisation of the S_NAr reaction, where *ortho* and *para* yields are selected as the objective for the optimisation. 83

Figure 24. Results of the five-parameter mixed variable optimisation on the S_NAr reaction, where an initial 25 experiment LHC was conducted with a further 74 additional iterations identifying 20 non-dominated solutions. Solvent shapes represent: ● – DMF, – NMP, ■ – EtOH, + – MeCN, × – DMAc. 85

Figure 25. Parallel coordinate plots for NMP and DMAc for the Pareto optimal solutions of each discrete variable highlighting the required interactions between the four continuous variables for each non-dominated solution. The error for % yield of both *ortho* and *para* was calculated using 10 repeat experiments from the optimisation, where an error of ±4% was observed for *ortho* % yields and an error of ±3% was observed for *para* yields. 87

Figure 26. Parallel coordinate plots for DMF an MeCN for each Pareto optimal solution produced by these solvents. The error for % yield of both *ortho* and *para* was calculated using 10 repeat experiments from the optimisation, where an error of ±4% was observed for *ortho* % yields and an error of ±3% was observed for *para* yields. 88

Figure 27. Hypervolume plot for the S_NAr reaction that was monitored throughout the optimisation procedure. After 60 experiments monitoring enabled the termination at 100 experiments once a plateau was achieved. 89

Figure 28. Parallel coordinates plot for the simulated Pareto front using polarity index as an input continuous variable. Residence time was maintained at 2 minutes and thus excluded from the figure. Simulation and figures were produced by Jamie Manson. 90

Figure 29. 12 principles of green chemistry with highlighted areas addressed by these types of mixed variable optimisations presented. 95

Figure 30. (A) Schematic for the Suzuki-Miyaura coupling reaction optimised with respect to mixed variables on a micro-fluidic platform. 97

Figure 31. Flow chart for the MVMOO self-optimisation procedure with inclusion of details for the timers throughout the process. 102

Figure 32. Exploration of different process metric outcomes using a 12 initial reaction data set, evaluating productivity metrics. 103

Figure 33. Schematic used for the optimisation of the Sonogashira cross-coupling case study. Where L1 is DavePhos, L2 is XPhos, L3 is CyJohnPhos, L4 is SPhos and L5 is TPP. The catalyst is $Pd(OAc)_2$, CuI and pyrrolodine base were included in each of these stock solutions. Additionally, R6 represents 2-bromo-4-(trifluoromethyl) benzonitrile (3.4), I.S is the internal standard 1,3,5-trifluoromethoxybenzene and R7 represents 3,3-dimethylbutyne (3.5) 106

Figure 34. Results of the four-parameter mixed variable multi-objective optimisation of the Sonogashira cross-coupling reaction. An initial 25 experiments were completed from LHC sampling with an additional 44 iterations direct by MVMOO, 12 of which formed a Pareto front highlighting the trade-off in STY and RME. Ligand shapes represent: ● – DavePhos, – XPhos, ■ – CyJohnPhos, × – SPhos, + – TPP. 107

Figure 35. Plot of the calculated yield from RME values vs experiment number, highlighting the outstanding yield achieved of 90% using the conditions corresponding to the optimum RME using TPP. Ligand shapes represent: ● – DavePhos, – XPhos, ■ – CyJohnPhos, × – SPhos, + – TPP. 108

Figure 36. Results from the Sonogashira coupling optimisation with individual process metrics against experiment number to highlight metrics enhancements over the optimisation process. Where ■ – Initial, ■ – Algorithm and ■ – Optimum. 109

Figure 37. Parallel coordinates plot showing the interactions between the variables for the Pareto optimal solution for the STY that was achieved by SPhos. 110

Figure 38. Parallel coordinate plot highlighting the interactions between the variables for each non-dominated solution of the Sonogashira optimisation. Each line represents a single Pareto optimal solution for TPP ligand points. Line colour is scaled in relation to STY weighting to aid in visualisation (high STY/low RME = –, low STY/high RME = –). 111

Figure 39. 3-D plot for the Pareto front solutions with respect to RME, where ● represents TPP based solutions and ■ represents the singular SPhos solution. 112

Figure 40. 3-D plot for the Pareto front solutions with respect to STY, where ● represents TPP based solutions and ■ represents the singular SPhos solution. 113

Figure 41. Calculated hypervolume vs experiment number for the Sonogashira optimisation that was used to determine the stopping criteria for this workflow. 114

Figure 42. Flow schematic for the synthesis of 1,2,4-triazoles using continuous telescoped flow techniques.[151] 119

Figure 43. Multipoint sampling technique used for telescoped optimisation with daisy-chained Vici sampling configuration where a single HPLC analysed both steps in the reaction. Configuration based on the work completed by Clayton *et al.*[20] 128

Figure 44. Flow schematic used for the telescoped optimisation case study. Where R8 is bromobenzene (4.5), L1 is dppm, L2 is dppe, L3 is dppp, L4 is DavePhos and L5 is TPP. The catalyst $Pd(OAc)_2$, NEt_3 base and internal standard methyl p-tolyl sulfone were included in each of these discrete variable stock solutions. Additionally, R9 represents ethylene glycol vinyl ether (4.6), and the solvent is ethylene glycol:MeCN (2:1 ratio). 131

Figure 45. Results of the five-parameter mixed variable multi-objective optimisation of the telescoped Heck-hydrolysis reaction. For each ligand, 5 initial LHC points were afforded. The MVMOO algorithm then iteratively ran an additional 41 experiments, 8 of which were identified as non-dominated solutions highlighting the trade-off between STY_{Diox} (4.8) and acetophenone (4.9) overall yield. Ligand shapes represent: ● – dppm, – dppe, ■ – dppp, × – DavePhos, + – TPP. 133

Figure 46. Results from the Heck-hydrolysis optimisation, with STY_{Diox} vs Experiment number plotted to highlight improvements made to this metric over the process. Where ■ – Initial, ■ – Algorithm and ■ – Optimum. 136

Figure 47. Results from the Heck-hydrolysis optimisation, with RME vs Experiment number plotted to highlight improvements made the RME metric over the process. Where ■ – Initial, ■ – Algorithm and ■ – Optimum. 136

Figure 48. Parallel coordinate plot highlighting the interactions between continuous variables and objective outcomes for each Pareto optimal solution. 138

Figure 49. Hypervolume vs. Experiment number plot for the telescoped Heck-hydrolysis reaction that was monitored throughout the optimisation process. 139

Figure 50. Photo of the automated mixed variable flow reactor used for Chapters 2, 3 and 4..... 150

Figure 51. Screenshot of the concentration profiles used within the simulated code for the S_NAr reaction. 155

Figure 52. Screenshot of the ODE Solver code required for the simulation of concentration profiles, returning the TimeData and Concentration Data, used to calculate the % yield results for each regioisomer in the simulation..... 156

Figure 53. Example HPLC chromatogram for the S_NAr optimisation. Retention times (min): bis 5 = 1.14; para 4 = 1.53; 2,4-difluoronitrobenzene 1 = 1.69; ortho 3 = 1.95; biphenyl (IS) = 3.86 .. 166

Figure 54. Flow schematic of the optimisation platform used for this chapter's work. R1 represents 2,4-difluoronitrobenzene 2.1 and R2 is morpholine 2.2. Biphenyl (internal standard) was included in all the R1 stock solutions. S1 is DMF, S2 is NMP, S3 is EtOH, S4 is MeCN and S5 is DMAc. 166

Figure 55. Flow schematic used for the Sonogashira optimisation. R6 represents 2-bromo-4-(trifluoromethyl) benzonitrile 3.4, I.S is the internal standard 1,3,5-trimethoxybenzene and R7 represents 3,3-dimethylbutyne 3.5. L1 is DavePhos, L2 is XPhos, L3 is CyJohnPhos, L4 is SPhos and L5 is TPP. The catalyst $Pd(OAc)_2$, CuI and pyrrolidine base was included in each of these stock solutions. 176

Figure 56. Example HPLC chromatogram for the Sonogashira reaction. Retention times (min): 2-chloro-4-(trifluoromethyl)benzonitrile 3.4 = 3.37 (230 nm); 2-(3,3-dimethylbut-1-yn-1-yl)-4-(trifluoromethyl)benzonitrile 3.6 = 4.44 (254 nm); 1,3,5-trimethoxybenzene (IS) = 2.17 (210 nm). 177

Figure 57. Example HPLC chromatogram at 230 nm for the Heck-hydrolysis reaction. (a) Heck-intramolecular cyclisation (first step); (b) Hydrolysis deprotection (second step). Retention times (min): Bromobenzene 4.5 = 5.74 and 13.05 min; Dioxolane 4.8 = 4.09 and 11.93 min; acetophenone 4.9 = 2.92 and 10.19 min and methyl p-tolyl sulfone (IS) = 2.44 and 9.71 min. 187

Figure 58. Flow schematic used for the Heck-hydrolysis telescoped optimisation. R8 represents bromobenzene 4.5 and R9 represents ethylene glycol vinyl ether 4.6. L1 is dppm, L2 is dppe, L3 is dppp, L4 is DavePhos and L5 is TPP. The catalyst $Pd(OAc)_2$, triethylamine base and internal standard methyl *p*-tolyl sulfone were included in each of these stock solutions. 188

List of Schemes

Scheme 1. Example of the selective *in-situ* generation of acyl-substituted aryllithium (1.4) and reaction with phenyl isocyanate electrophile. 27

Scheme 2. Heck reaction between 4-chlorobenzotrifluoride (1.8) and 2,3-dihydrofuran (1.9) to form the mono-(1.9) and bi-substituted products used in the Nelder-Mead Simplex optimisation. 43

Scheme 3. Sonogashira coupling reaction optimised using a TSEMO algorithm between 3,5-dibromopyridine (1.12) and 1-hexyne (1.13) to form the desired mono-substituted alkyne (1.14) and undesired bis-alkyne (1.15). 54

Scheme 4. Suzuki-Miyaura cross-coupling reaction between *(E)*-vinyl sulfonate (1.16) and aryl boronic acid (1.17) used for the mixed-variable optimisation where maximisation of the *(E)*-desired product (1.18) was priority. 58

Scheme 5. S_NAr reaction between 2,4-difluoronitrobenzene (2.1) and morpholine (2.2) for the formation of the ortho- (2.3), para- (2.4) and bis- (2.5) regioisomers. 71

Scheme 6. Resonance stabilisation of the *ortho* regioisomer through a Meisenheimer complex. 72

Scheme 7. Kinetic reaction pathway for the formation of all regioisomers in the S_NAr pathway. 74

Scheme 8. Catalytic cycle for the Sonogashira cross-coupling reaction between an aryl or vinyl halide and terminal alkyne, facilitated by a palladium catalyst and copper (I) co-catalyst in a side catalytic cycle. 99

Scheme 9. Sonogashira cross-coupling reaction between aryl bromide 3.4 and terminal alkyne 3.5 to form aryl alkyne 3.6. Includes optimisation parameters. 100

Scheme 10. Catalytic cycle for the Heck reaction between an aryl or vinyl halide and alkene, facilitated by a palladium (0) catalyst which is regenerated via the addition of a base in the final step of the cycle. 123

Scheme 11. Neutral and Cationic pathways for alkene insertion within the Heck catalytic cycle for the direction of alpha and beta products.[157] 124

Scheme 12. Heck reaction between bromobenzene (4.5) and ethylene glycol vinyl ether (4.6) to form the alpha substituted product (4.7) followed by intramolecular cyclisation to form the dioxolane product (4.8). Subsequently, the hydrolysis deprotection reaction to form the final acetophenone product (4.9). 126

Scheme 13. Formation of the desired alpha- and undesired beta-products via the respective cationic and neutral pathways in the Heck reaction. 134

Scheme 14. Equilibrium for the formation of the cationic palladium(II)-alkene intermediate promoted by the presence of hydrogen-bond donation by ethylene glycol. 134

List of Tables

Table 1. Overview summary of examples utilising local optimisation algorithms on chemical systems. 44

Table 2. Comparison of Environmental and Economic metric objectives for evaluating chemical processes. 64

Table 3. Boundaries for continuous variables in the S_NAr simulation with morpholine. 76

Table 4. Chemical descriptor values of common solvent polarity metrics.[112] 82

Table 5. Boundaries for continuous variables in the telescoped Heck-hydrolysis optimisation. 130

Table 6. Parameter boundaries for the four-variable single-objective simulated self-optimisation of the S_NAr reaction using pyrrolidine and morpholine in separate studies for comparison. 156

Table 7. List of operating conditions and results from the simulated SNOBFIT optimisation of the S_NAr reaction for pyrrolidine with *ortho* yield as the objective. The optimal yield and conditions are highlighted in green. 157

Table 8. List of operating conditions and results from the simulated SNOBFIT optimisation of the S_NAr reaction for pyrrolidine with *para* yield as the objective. The optimal yield and conditions are highlighted in green. 159

Table 9. List of operating conditions and results from the simulated SNOBFIT optimisation of the S_NAr reaction for morpholine with *ortho* yield as the objective. The optimal yield and conditions are highlighted in green. 161

Table 10. List of operating conditions and results from the simulated SNOBFIT optimisation of the S_NAr reaction for morpholine with *para* yield as the objective. The optimal yield and conditions are highlighted in green. 162

Table 11. List of reservoir solutions for the S_NAr optimisation. Biphenyl was used as the internal standard for the optimisation process. 165

Table 12. Parameter boundaries for the five-variable experimental multi-objective self-optimisation of the S_NAr reaction using the MVMOO algorithm. 167

Table 13. List of operating conditions and results from the self-optimisation of the S_NAr reaction. The first 25 experiments were completed as a LHC design. Pareto optimal points are highlighted in green. 167

Table 14. List of Pareto front experiments in order of optimum *ortho*-2.3 yield % to optimum *para*-2.4 yield %. 171

Table 15. List of reservoir solutions for the Sonogashira optimisation. 1,3,5-trimethoxybenzene was used as the internal standard for the optimisation process. 174

Table 16. Parameter boundaries for the four-variable self-optimisation of the Sonogashira coupling reaction. 178

Table 17. List of operating conditions and results from the self-optimisation of the Sonogashira coupling reaction. The first 25 experiments were completed as a LHC design. Pareto optimal points are highlighted in green. 178

Table 18. List of Pareto front experiments in order of optimum RME to optimum STY. 181

Table 19. List of reservoir solutions for the Heck-hydrolysis telescoped optimisation. Methyl p-tolyl sulfone was used as the internal standard for the optimisation process. 185

Table 20. List of operating conditions and results from the self-optimisation of the mixed variable multi-objective Heck-hydrolysis telescoped reaction. The first 25 experiments were completed as a LHC design. Pareto optimal points are highlighted in green. 188

Table 21. List of Pareto front experiments in order of optimum STY_{Diox} to optimum overall yield. 191

List of Abbreviations

A	Pre-exponential factor
API	Active pharmaceutical ingredient
Boc	*tert*-Butyloxycarbonyl
C	Concentration
cm	Centimetre
CSTR	Continuously stirred tank reactor
D	Diameter
Da	Damköhler number
DCM	Dichloromethane
DMAc	Dimethylacetamide
DMC	Dimethyl carbonate
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
d	Doublet
DoE	Design of Experiments
Diox	Dioxolane
dppm	bis(Diphenylphosphino)methane
dppe	bis(Diphenylphosphino)ethane
dppp	1,3-bis(Diphenylphosphino)propane
E_a	Activation energy
EIM	Expected improvement matrix
Eq.	Equivalents
Equiv.	Equivalents
ESI^+	Positive mode electron spray ionisation
Et	Ethyl

EtOH	Ethanol
FFD	Full factorial design
g	Gram
GP	Gaussian process
HPLC	High performance liquid chromatography
HTE	High-throughput experimentation
HTS	High-throughput screening
Hz	Hertz
ID	Inner diameter
IS	Internal standard
J	Coupling constant
k	Rate constant
LHC	Latin hypercube
m	multiplet
Me	Methyl
MeCN	Acetonitrile
MHz	Megahertz
min	minute
MINLP	Mixed integer nonlinear programming
mL	Millilitre
mol	Moles
MVMOO	Mixed variable multi-objective optimisation
MW	Molecular weight
MΩ	Megaohm
n	Number of moles
NMP	*N*-Methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance

NMSIM	Nelder-Mead simplex
OD	Outer diameter
ODE	Ordinary differential equation
OVAT	One-variable-at-a-time
P	Pressure
Ph	Phenyl
PhMe	Toluene
PSI	Pounds per square inch
PTFE	Polytetrafluoroethylene
R	Gas constant
RAFT	Reversible addition-fragmentation chain-transfer
Re	Reynolds number
R&D	Research and development
RME	Reaction mass efficiency
s	Second or singlet
sc	Supercritical
SM	Starting material
SMSIM	Super modified simplex
S_NAr	Nucleophilic aromatic substitution
SNOBFIT	Stable noisy optimisation by branch and fit
STY	Space-time yield
T	Temperature
t	Time or triplet
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TMS	Trimethylsilane
TOF	Turn over frequency

TPP	Triphenylphosphine
tR	Residence time
t_{res}	Residence time
TSEMO	Thompson sampling efficient multi-objective
μ	Viscosity
µL	Microlitre
uPLC	Ultra performance liquid chromatography
V	Volume

Chapter 1. Introduction

1.1 Fundamentals of Continuous Flow

Continuous flow chemistry, often described as an enabling technology, represents an alternative technique for executing chemical transformations and reactions compared to conventional batch techniques. The essence of flow chemistry lies in its name – it involves the seamless passage of reagents as a continuous stream through a network of reactors until the desired transformation is achieved. In contrast, batch reactors typically unfold within enclosed containers like round-bottom flasks. Flow chemistry, however, unfolds within chemically resistant tubing, allowing the reaction mixture to continuously flow. This difference in apparatus sets the stage for some fundamental differences between the two approaches.

One crucial disparity is how the chemical reactions progress. In flow, the progression of a reaction hinges on the distance travelled by the reagents within the tubing. As a result, the composition of the reaction mixture changes gradually as it moves along this path in the reactor. During this sequence, the reagents in the system will encounter different reaction conditions e.g. the temperature of reactors, the pressure of the system and mixing, ultimately leading to changes in the reagent concentrations, Figure 1. On the other hand, batch chemistry relies on the quantity of time as the primary driver of composition changes. The longer the mixture is stirred at the set reaction conditions, the more it transforms.

This distinction gives rise to a crucial concept in flow chemistry known as “steady-state”. It refers to conditions where the chemical species within the mixture remain stable over time. In simpler terms, the components within the reaction mixture reach a point where their concentrations remain relatively constant. This state-state concept is an essential asset to this technique, offering researchers a high level of control over reaction variables to drive reaction pathways to the desired outcomes consistently.

The ”steady-state” concept in flow chemistry is the notion of achieving stability in concentrations of chemical species within the reaction mixture over time. It enables researchers to maintain consistent reaction conditions, which is crucial for controlling and optimising chemical reactions in flow systems.

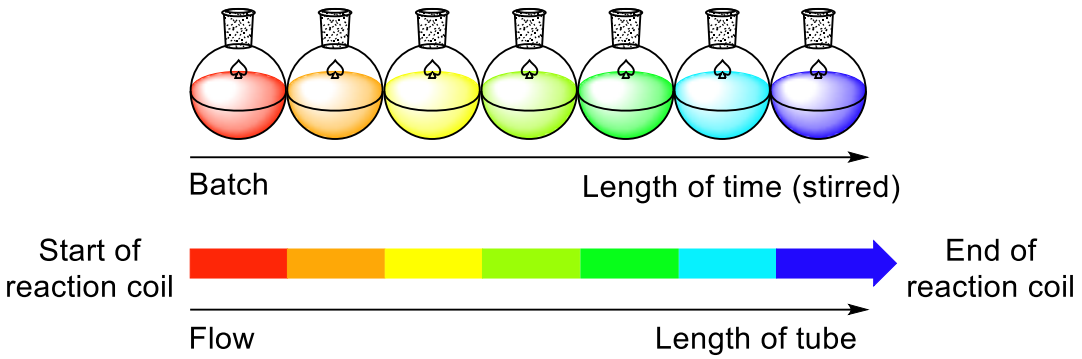


Figure 1. Comparison of reaction coordinates between flow technologies and conventional batch methodologies

As previously mentioned, the continuous nature of flow technologies demands entirely different equipment compared to traditional batch processes. To carry out continuous flow chemical reactions, a specialised apparatus is essential. The increasing popularity of this technology has led to advancements in the equipment becoming readily available, ranging from basic to highly specialised setups. This variety ensures that researchers have access to the necessary tools for conducting continuous flow experiments.

In recent decades, continuous flow processes have gained significant popularity in both the pharmaceutical and fine chemical industries.[1,2] This heightened interest has driven substantial enhancements and refinements in this technique over the years. Moreover, it has been integrated with other chemical methodologies such as photochemistry, microwave chemistry and electrochemistry, highlighting the versatility and capabilities of continuous flow. This feature has made it a valuable tool for chemists in both research and industry to exploit.[3] Additionally, ongoing improvements in various aspects of continuous flow components, such as reactor design, tubing types and sizes have contributed to enhancing the efficacy of this enabling technology. These modifications reinforce the advantages that flow chemistry can offer when integrated with other chemical processes.

One significant advancement is the design of different reactor types which have been further implemented to enhance the capabilities of flow chemistry within synthetic and large-scale production laboratories. The development of new types of reactors plays a prominent key role in the

effectiveness of the technique as the reactor is where the main chemical reactions take place. These provide unique advantages over one another, such as seen with the use of packed bed reactors, which allows for a solid such as a catalyst to be loaded into a cartridge and placed such that the reaction mixture is streamed through the cartridge.[4] Implementing this approach enables the reactants within the mixture to interact with the solid substances immobilised within them. These solid materials serve various functions in a reaction, such as loaded catalysts, bases, or drying agents among others which might not be possible to dissolve in solution. The inclusion of these solids in a reactor bed addresses one of the main challenges that flow chemistry poses, which is the integration of solid components into reaction pathways. Enabling the access to suspended solid substances within a reactor bed, can unlock large libraries of bio- or catalytic chemistry that might otherwise remain inhibited.

Alternatively, another innovative solid-liquid hybrid reactor design is the continuously stirred-tank reactor (CSTR), offering an alternative to conventional reactor coils, for managing solid or slurry-type reaction mixtures.[5] A typical CSTR consists of a magnetic stir bar placed within a small chamber, allowing the solid-liquid reaction medium to flow in and out of the heated reactor compartment. These reactors are often arranged in a cascade or series, with multiple reactor beds positioned sequentially to ensure complete conversion and the desired product formation. This design has found insightful applications for photo-flow chemical optimisation, where Manson *et al*. strategically positioned a UV lamp above the top of the reactor to facilitate a photochemical reaction, Figure 2. This arrangement enabled a case study for the aerobic oxidation of tetralin (**1.1**) to tetralone (**1.2**) utilising benzophenone as a photo synthesiser for this formation, Figure 2.[6] This was then further optimised using a hybrid-based algorithm to achieve yields of up to 65%.

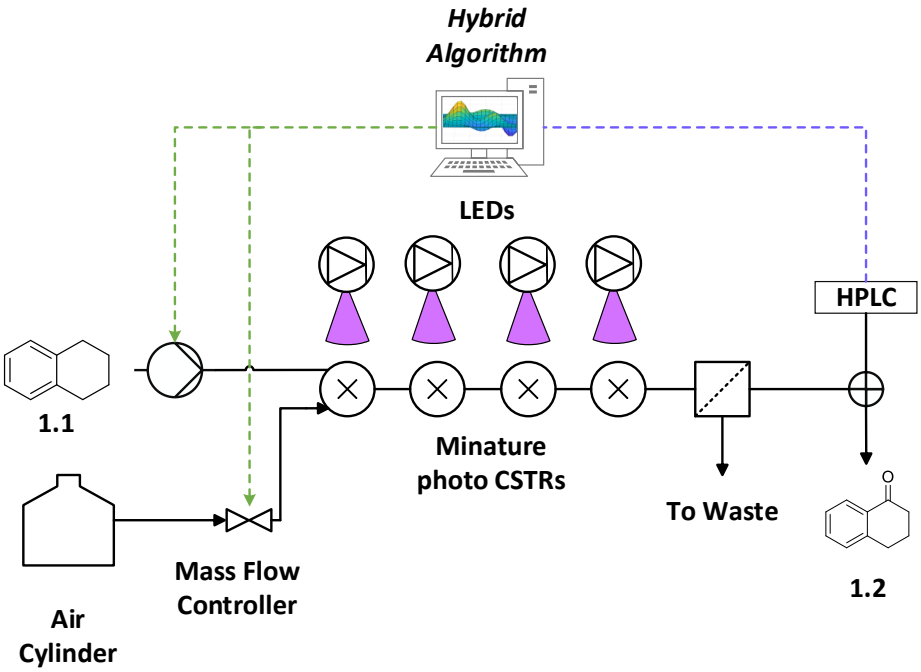


Figure 2. Photo-flow platform used in the optimisation utilising a four miniature CSTR cascade for the aerobic oxidation of tetralone from tetralin with benzophenone as a photo synthesiser.

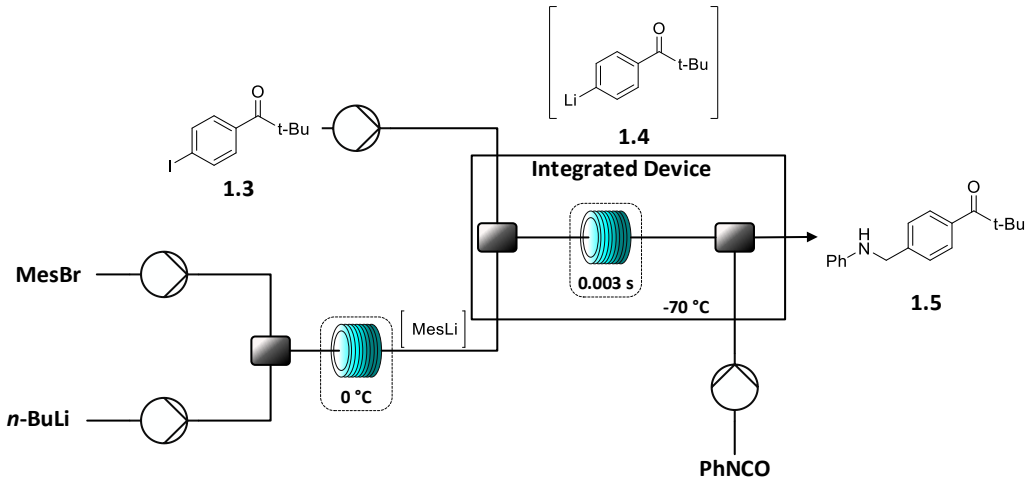
1.2 Advantages of Flow Chemistry

Reactor design represents a notable advantage that continuous flow processes offer when integrated with other enabling technologies, enhancing reliability and scalability in chemical synthesis. However, flow methodologies offer additional substantial benefits over traditional batch techniques for synthetic processes. This approach enables the implementation of potentially problematic batch reactions with a greater level of safety, whilst also increasing the overall efficiency of the procedure.

1.2.1 Containment of Hazardous Chemicals

The advantages of flow chemistry extend to enhancing safety in multiple ways. Firstly, the small volumes inherent in flow systems mean that only a limited quantity of the mixture is exposed to reaction conditions such that the generation of hazardous intermediates at any given time is significantly reduced. This fundamentally reduces the potential for exposure to quantities of harmful or toxic substances, thereby enhancing the overall safety of this technique. Moreover, this facilitates the design of flow processes that exploit *in-situ* generation of hazardous intermediates to be utilised

downstream and unlock reaction pathways that would have otherwise remained inaccessible. This enhanced safety of hazardous intermediate generation was manipulated by Kim *et al.* where the development of a microreactor to facilitate extremely short reaction times enabled the implementation of a reaction between an organolithium species and a carbonyl, Scheme 1.[7] Traditionally, the reaction between carbonyl groups and organolithium species is extremely reactive, typically requiring functional group protection to moderate the reactivity and selectivity.[8] Implementation of an integrated microreactor facilitated residence times of 0.003 seconds to be achieved, this enabled the selective *in-situ* generation of acyl-substituted aryl lithiums that could be immediately quenched with an electrophile. Using this methodology, lithiation of *p*-dimethylpropanoyliodobenzene (**1.3**) was generated *in situ* via an iodine-lithium exchange process with mesyllithium to form the corresponding acetylphenyllithium (**1.4**). This short-lived intermediate was directly reacted with a phenyl isocyanate electrophile to form the desired product (**1.5**) achieving an 87% overall yield.



Scheme 1. Example of the selective *in-situ* generation of acyl-substituted aryllithium (**1.4**) and reaction with phenyl isocyanate electrophile.

Additionally, the continuous nature of a flow process itself contributes to safer experiment procedures. The seamless flow of reaction mixture minimises the risk of accumulating explosive intermediate within the system. In contrast, batch reactions can pose a substantial risk if explosive intermediates form, as a significant proportion of the flask might contain these

volatile compounds, increasing the potential for accidents. Furthermore, the absence of a headspace in the flow tubing serves as an added safety precaution against the accumulation of volatile explosive intermediates.[9] Flow chemistry proves especially advantageous when handling highly exothermic reactions that have the potential to trigger thermal runaway which would result in severe accidents when conducted in batch systems. In flow systems, the risk of thermal runaway can be effectively mitigated through the improved heat dissipation facilitated by the small volumes of liquid contained within the tubes. This heightened control over heat not only ensures the safer execution of such reactions but also grants greater control over the reaction process.[10] Moreover, the enhanced heat dissipation attributes of flow setups eliminate the formation of potential “hot spots”, consequently reducing the chance of side reaction or product overreaction/thermal decomposition. This additional safety feature not only averts accidents but also enhances the reliability and reproducibility of flow reactions.

1.2.2 Mixing and Mass Transfer

In a continuous flow system, reaction mixtures are continuously introduced to one another, allowing for the observation of mixing effects between the incoming streams. This contrasts with traditional batch reactions, where all reactants are combined at once and allowed to react in a partially static environment. The efficiency of mixing in flow systems is a critical factor that directly impacts the reaction conversion and performance which can be assessed by considering different mixing modes that occur as streams combine within the flow tubing. One of the key parameters for describing these mixing patterns is the Reynolds number (Re), a dimensionless quantity widely employed in fluid dynamics to predict and characterise flow behaviours in such systems. The Reynolds number can be calculated according to [Eq (1)], where ρ (density, kgm^{-3}), v (linear flow rate, m/s), D (diameter, m) and μ (viscosity, $Pa \cdot s$) are the respective parameters:

$$\mathrm{Re} = \frac{\rho v D}{\mu} \qquad (\mathbf{1})$$

When the Reynolds number of a flow system drops below approximately 2000, the flow is classified as laminar and when the number is above 3000, the flow is characterised as turbulent.[11] In the turbulent flow regime, the dominant forces at play are the disordered changes in the velocity resulting in highly effective mixing among the fluid layers within the system. In contrast, for the laminar flow regime, the dominant forces are the viscous forces, and turbulence is virtually absent. Consequently, due to no disruption in the parallel fluid layers, mixing predominantly relies on the rate of diffusion taking place along the longitudinal fluid interface.[12] This type of mixing is often described as passive, as it occurs naturally due to the slow diffusion of molecules and is driven by concentration gradients within the system. Conventional batch reactors exhibit distinct mixing patterns, with turbulent regions near the stirring mechanism, creating rapid changes in velocity, and laminar regions closer to the vessel walls.[13]

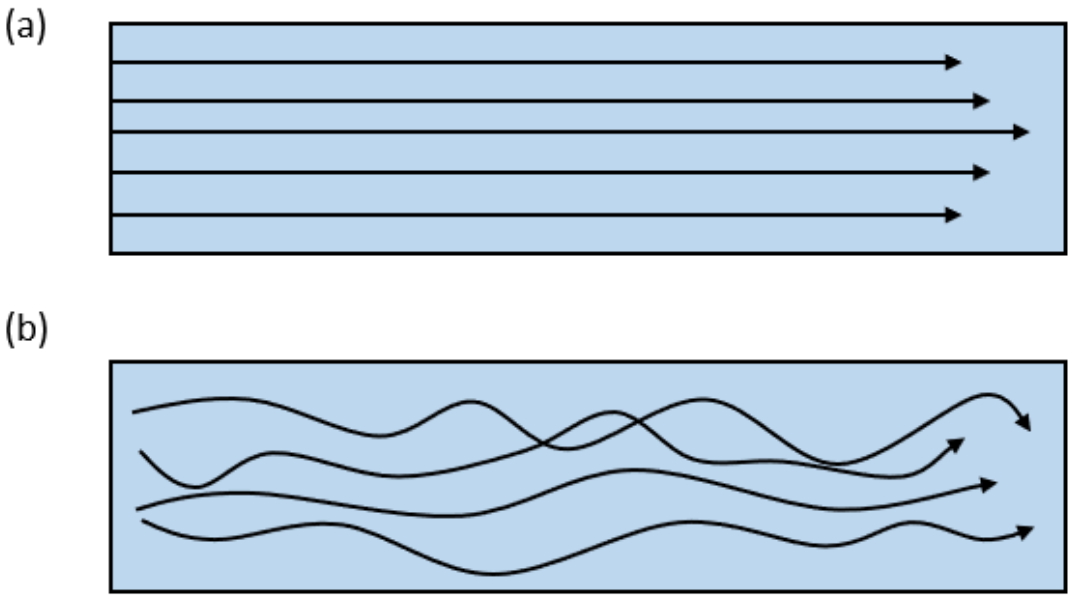


Figure 3. (a) Example of Laminar flow in tubing where there is no disruption between parallel fluid layers. (b) Turbulent flow where there is a presence of rapid changes in velocity.

On the other hand, standard laboratory-scale flow reactors typically exhibit laminar flow patterns. This tendency is a consequence of the combination of low flow rate operating conditions and the small dimensions characteristic of these reactors. When compared to batch vessels, tubular flow reactors exhibit a substantially higher surface area-to-volume ratio.

Consequently, this results in an increased rate of diffusion and subsequently an enhanced rate of mixing.

The mixing effects play a pivotal role in mass transfer within a flow system, and their influence can be elucidated through the use of the Damköhler number (Da), which defines the rate of reaction to the rate of diffusion within a system [Eq (**2**(**3**)].[14] Where *k* is the rate constant of the reaction, C_0 is the initial concentration, *n* is the order of reaction, d_t is the diameter of the tubing and *D* represents the ratio of the rate of reaction to the rate of diffusion.

$$\mathrm{Da} = \frac{\mathrm{k}\mathrm{C}_0^{\mathrm{n}-1}\mathrm{d}_\mathrm{t}^2}{4\mathrm{D}} \qquad (\mathbf{2})$$

$$\mathrm{D} = \frac{\text{rate of reaction}}{\text{rate of diffusion}} \qquad (\mathbf{3})$$

When the Da > 1, it signifies that the rate of reaction surpasses that of the mass transfer of the system. This scenario can result in the accumulation of a concentration gradient, which, in turn, can have detrimental effects on the reactor's efficiency. Additionally, it may lead to an increased production of undesirable byproducts within the reaction process.[9,14] In the case of Da < 1, it means that the rate of mass transfer by diffusion is greater than the rate of reaction such that an essentially homogenous reaction mixture is achieved. To address challenges stemming from limited mass transfer when Da > 1, flow reactors have demonstrated their efficiency in eliminating concentration gradients, owing to their aforementioned enhanced mixing capabilities. Furthermore, the development of micromixers and microreactors in flow systems has successfully proven to circumvent potential complications arising from limited rates of diffusion, especially in cases involving rapid reactions.[15,16] Modelling of the Damköhler number has been performed numerous times within literature, classically through the use simulated computational studies, where values for Da are typical reported between $10^{-2} - 10^{3}$.[17,18]

1.2.3 Unlocked Temperatures and Pressures

In comparison to traditional batch methods, the utilisation of a flow setup significantly amplifies the efficiency of heat transfer processes. This

remarkable enhancement can be attributed to the substantially increased surface area-to-volume ratios intrinsic in the system. As previously highlighted, this heightened heat dissipation capability inherent in flow setups not only facilitates the safe execution of highly exothermic reactions but also frequently removes the necessity for extremely harsh cryogenic conditions. Fukuyama *et al.* emphasised this possibility, where they were able to perform flow experiments for the diisobutylaluminium hydride reduction of an ester to aldehyde at temperatures elevated from -70 °C in batch to -50 °C in an optimised flow system.[19] Furthermore, they went on to emphasise their achievement in conducting an *n*-butyllithium-mediated coupling of the aldehyde product with a sulfone fragment at substantially higher temperatures and better conversions, specifically 10 °C, within a flow system, as opposed to the demanding cryogenic conditions of -70° C necessary in batch procedures. This accomplishment presents the potential for a substantial reduction in the need for cryogenic temperatures for these transformations when performed under continuous flow conditions.

One significant advantage of flow platforms lies in the capacity to apply high pressure to the system using a back-pressure regulator. This ability to manipulate higher temperatures and pressure safely, such that it enables solvents to be heated to temperatures far exceeding their usual boiling points. Consequently, this allows for the execution of chemical reactions under optimal conditions, employing the desired solvents, without the necessity of costly high-pressure batch vessels. The elevated temperatures at which experiments can be conducted translates to accelerated reaction rates, as defined by the Arrhenius equation, leading to short reaction durations and increased production rates, for instances where selectivity is not impacted by higher temperatures. These faster reaction rates can aid in the reduction of material required to carry out these processes, presenting these types of methodologies as overall greener, provided that the energy consumption to reach the elevated temperatures is not excessive.

1.2.4 Telescoped Synthesis

When striving to synthesise complex natural compounds in batch processes, challenges often arise related to the manner in which the

technique is carried out. Firstly, traditional research adopted a “one-pot” synthesis approach, which involved the combination of various reagents in a singular vessel with attempts to directly utilise intermediates formed within a system for product generation. However, this approach comes with obvious disadvantages as it entails the combination of multiple reagents, potentially leading to undesirable side reactions or byproduct generation, hindering the reaction’s effectiveness and conversions. Additionally, for reaction pathways that include the generation of unstable intermediates, significant challenges arise in maintaining the reaction species concentration for further reaction without degradation occurring, which can further limit the overall process efficiency. The alternative to performing singular vessel synthesis for the multistep reaction is performing singular steps in a reaction pathway, followed by workups, purification, isolation, and quantification for further reaction. This is approached in an iterative loop with many active pharmaceutical ingredients (APIs) being traditionally synthesised in this manner. This methodology only leads to the generation of substantial amounts of waste through undesired workup media, purification solvent systems and analysis waste but also results in elongated production times due to each step of a reaction pathway having its own additional sets of processes required. Furthermore, when employing this workflow for large-scale synthesis of APIs, it exposes potential vulnerabilities. In the event of any disruption in the supply chain for production, it could result in substantial delays or shortages in the output, affecting both efficiency and cost-effectiveness.[20]

In contrast, flow chemistry provides a unique capability for telescoped synthesis, where in this approach, reactions are compartmentalised, allowing for precise control and sequential addition of reagents through each segment during the multistep pathway.[21] The modular nature of flow systems facilitates the ability to construct complex platforms with relative ease and adaptability, allowing for streamlined customisation and rapid scalability. The capacity to provide quench streams/reactors and in-line purification methods after each stage in a reaction addresses the difficulties encountered by traditional batch methodologies where intermediates present in further reactions can detrimentally affect the reaction outcomes. Furthermore, the capacity to quench, purify and quantify reagents through inline or online analysis all within

a singular, continuously flowing process highlights the significant potential that these types of workflows can pose when performed in a telescoped operation.[22]

Murray *et al.* highlighted the ability to apply these types of workflows for the total synthesis of Tamoxifen from simple and commercially available starting materials.[23] This innovative approach involves the development of a five-stage flow reactor system that incorporated premixing and quench coils to minimise the potential of unwanted downstream side reactions. Additionally, this platform provided the capability to access varying temperature zones along the pathway, enhancing the overall process efficiency. Utilising this telescoped system, a total synthesis of Tamoxifen was achieved with an impressive 85% overall yield. Remarkably, just 80 minutes of continuous pumping resulted in the production of 12.4 grams of the desired product, all of which highlight the significant capabilities of flow systems for performing multi-step reactions in a telescoped manner.

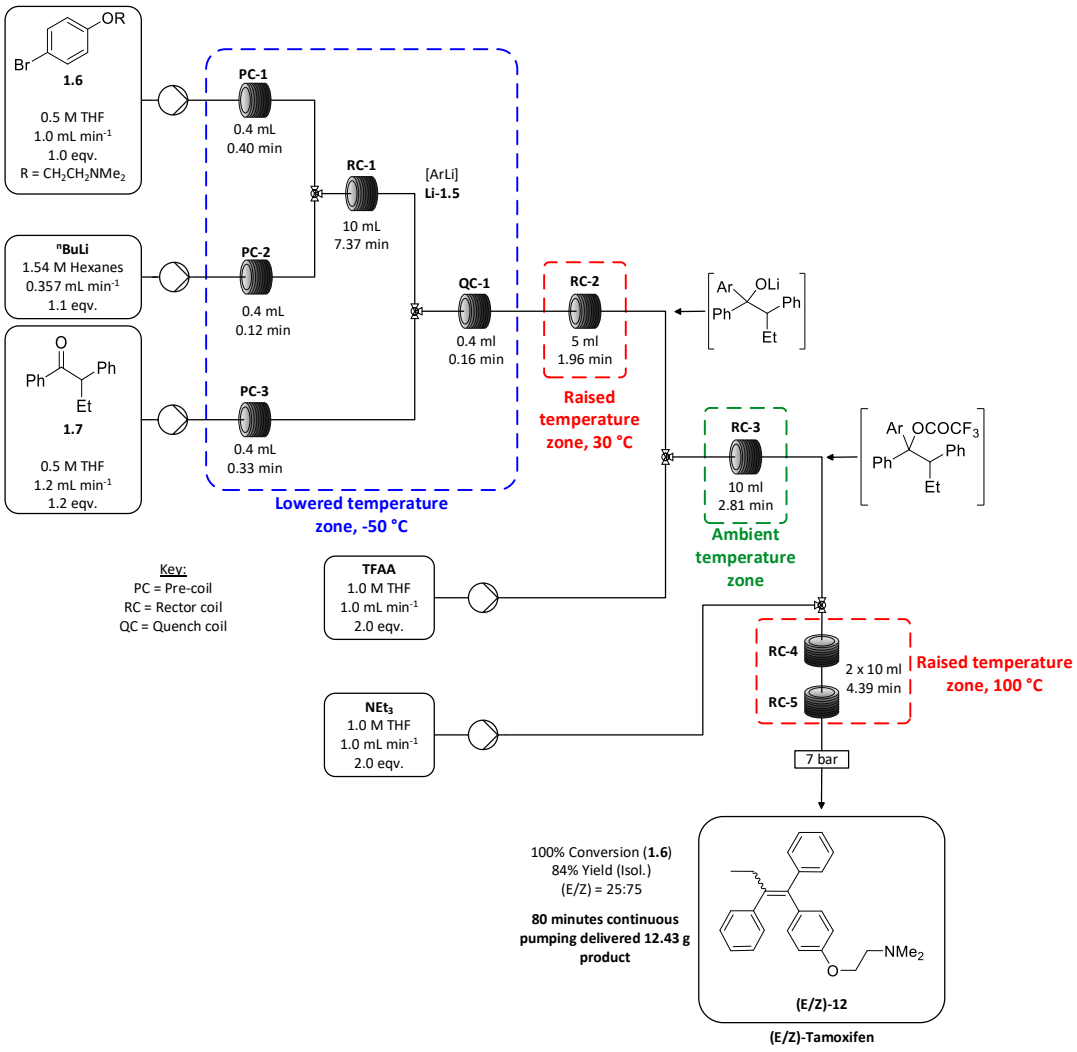


Figure 4. The telescoped flow platform employed in the synthesis of Tamoxifen is equipped with various temperature zones for reactions, in addition to pre-mixing and quenching coils.[23]

1.3 Automated Self-Optimisation

The introduction of advanced technologies in chemical synthesis has sparked significant interest in automating chemical processes, with a primary focus on reducing the time and labour required for process development.[24] Application of automated synthesis has been observed through both batch and flow experimentation where control of the platforms has enabled enhanced management of reaction pathways, thus minimising labour requirements. This growing interest has extended to the automated optimisation of chemical systems, leveraging statistical or algorithmic approaches to design or dictate experimental conditions for these platforms.[25]

Self-optimisations of chemical systems aim at maximising or minimising objective metrics via the manipulation of the reactions variables. A chemical systems' variables can be split into two categories: continuous and discrete. Continuous variables are parameters that can be measured on a continuous scale, for example temperature, concentrations, residence times. On the other hand, discrete variables are those that can only take on specific values, often representing categorical choices within the experimental setup, such as the choice of solvent, catalyst, ligand etc. Continuous variables allow for fine-tuning and adjustment within a range, facilitating precise control over reaction conditions and parameters. They are typically chosen based on their impact on reaction kinetics and objective metrics. Discrete variables introduce a dimension of choice and strategy into the optimisation process, where selection of the discrete variables such as the choice of reactants can significantly influence the outcomes of the reactions and the efficiency of the overall process. Therefore, careful consideration is given to choosing discrete variables that offer the most favourable conditions for the desired chemical transformations. Ultimately, the selection and manipulation of both continuous and discrete variables are guided by the principal goal of optimising desired outcomes, whether it is maximising yield, selectivity, efficiency, or other process metrics, in the most effective and efficient manner possible.

The inherent capabilities of flow systems, heavily facilitates the employment of optimisation algorithms onto them to enable reaction parameter control. To permit the use of algorithms in these automated flow platforms, computers have been integrated into the process in a closed feedback loop, whereby data from in/online analytical equipment is sent through to the computer in an iterative manner, Figure 5.[26–28]

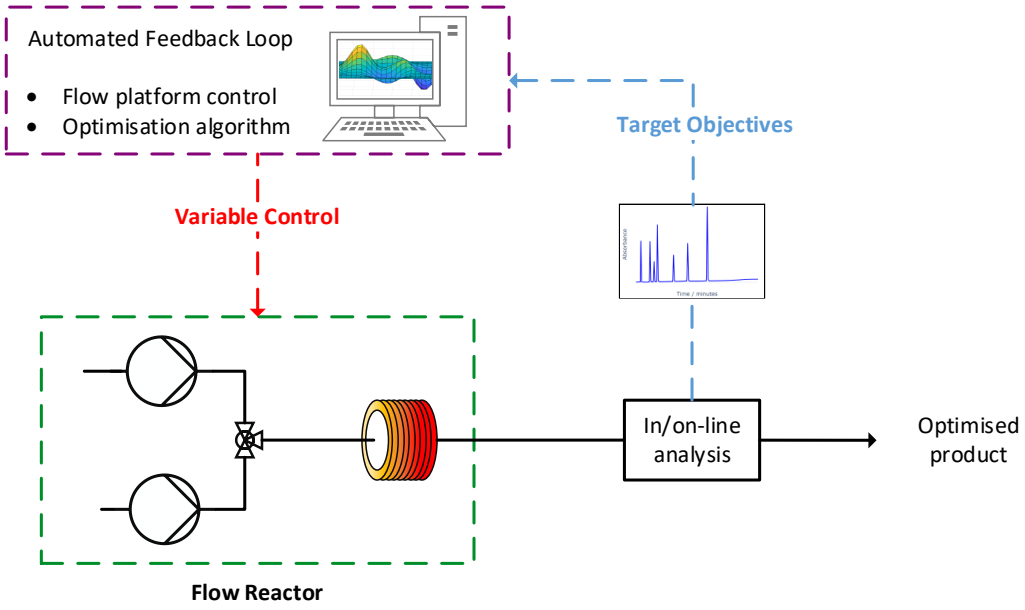


Figure 5. Illustrative example of a self-optimisation flow platform with an automated feed-back loop for system control and direction of conditions.

This process allows for the continuous analysis of data, which is then used by an algorithm to instruct the process, including the control of key reaction parameters such as temperature, equivalents, concentrations, and residence times. This synergy enables complete automated self-optimisation of chemical processes, eliminates the need for human intervention and significantly reduces labour requirements.[29] Furthermore, applying these feedback loop systems for intelligent optimisation of the experimental domain reduces the number of experiments required, subsequently, resulting in a more efficient, cost-effective, and environmentally sustainable approach for optimisation of chemical systems.

1.3.1 Local Optimisation

1.3.1.1 One-Variable-at-a-Time

For decades, traditional optimisations performed within synthetic chemistry have mostly employed a one-variable-at-a-time (OVAT) approach for the refinement of chemical processes. OVAT is a methodical approach, where each variable is sequentially changed per experiment, whilst the remaining parameters are maintained at a set constant or at their previously optimised value, to assess the impact on the reaction outcome. An example of a two-variable OVAT optimisation is shown in Figure 6, where the variable

x_1 is initially studied and changed along the line ***A*** to ***B*** whilst x_2 is maintained at a constant value. Upon identification of the best x_1 result x_b, x_2 is varied along the line ***C*** to ***D*** with the value from x_b maintained, leading to the identification of the optimum point, ★. Whilst the OVAT approach offers a simplified methodology for the potential optimisation of a chemical system, it has several major disadvantages: (i) Exploration of experimental space is significantly reduced, restricting information gained about the system. (ii) Variable-variable interactions are not considered, which can significantly affect the process outcomes. (iii) Optimisation of the system is highly dependent on the starting values assigned for the initialisation of the process. (iv) Highly time and resource-intensive, with iterations at designated intervals performed along each variable limits it can increase the number of experiments required, especially in systems with several variables under study.

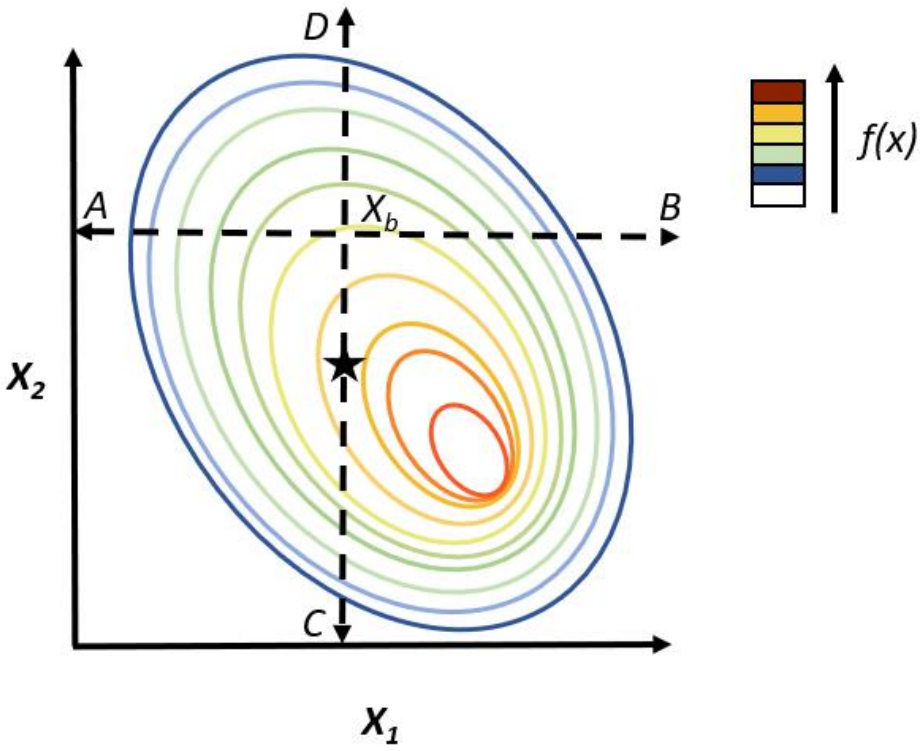


Figure 6. Example two variable OVAT optimisation leading to a false optimum to be observed. ★ represents the best results observed from the process.

Hence, the utilisation of the OVAT approach in chemical procedures can result in inefficient optimisation processes, potentially leading to the

optimum conditions for the system being missed. Moreover, when dealing with an increased number of variables, this issue is further intensified, reducing the likelihood of identifying the true optimum.[30] Additionally, this amplifies the demand for a greater number of experiments within the workflow, consequently leading to increased material consumption and time inefficiencies.

1.3.1.2 Design of Experiments

In the last decade, the literature has increasingly featured the application of Design of Experiments (DoE) for optimising chemical systems. DoE is a statistical-based systematic optimisation approach used to plan, conduct, and analyse experiments. The use of DoE at both industrial and academic scales has been prevalent within the literature with both sides employing it more frequently in their optimisation processes.[31,32] Unlike OVAT, DoE takes into account the synergistic effects that the variables being changed can have on one another. As previously highlighted, the assumption that the variables are independent within the OVAT approach can lead to a false optimum being achieved. However, for most chemical systems, the synergistic effects between the variables will be present and will most likely affect the outcome of the reaction such that the identification of the true optimum of a system requires the exploration of the design space.

Another advantage of DoE over OVAT is the facilitation of gaining knowledge of experimental information in a larger region of the design space. This means that the process understanding acquired about the response surface for the given factors and the interactions taking place is further increased. A response surface can be described as the relationship between experimental variables (e.g., temperature, pH etc.) and a response (e.g. Yield). This further improves the optimisation as it grants a greater region of the optimum to be explored for each factor and the relationship each factor can have with one another.[33]

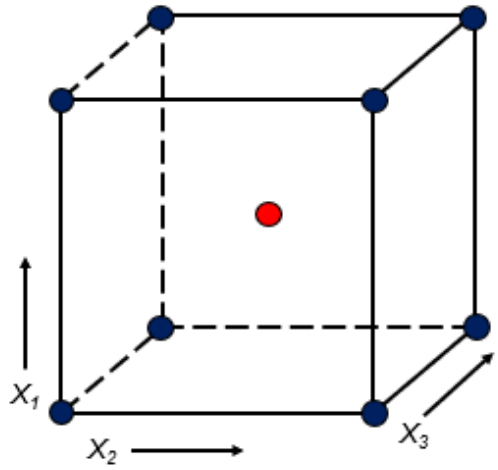
When planning to perform reactions by experimental design, there is a general set of rules which the user can follow to ensure that the experiments are designed and performed correctly. Not following the rules can lead to large numbers of unnecessary experiments being performed and as such large

amounts of wasted time and resources which can be expensive.[34] An example of some of the steps which have to be followed to plan a DoE is shown below.

1) Determine the aim for the experiment: Establish the objectives for optimisation.
2) Examine all factors that can potentially have an effect: All factors that can affect the outcome of the reaction are required to be identified.
3) Planning of the experiments: Once the factors are determined, their range limits (lower and upper bounds) can be set. Simultaneously, select and formulate the model being applied to the experimental process.
4) Execute the experiments: Perform the set experiments that have been generated by the design and record the results of each experiment.
5) Analyse the results from the experiments performed: Transform the collected data into a logical conclusion of the process, considering the interactions and effects of the factors on the process outcomes.

Full factorial designs (FFD) are the simplest form of design method that can be applied to DoE. They serve as a basic yet effective method for systematically exploring factors in the system that are expected to have a substantial effect on the response. FFDs can be designed to include several factors within the experiment, where the number of experiments required, *N,* can be calculated using [Eq (**4**)], where *n* is the number of levels, *k* is the number of factors being included and *m* is the number of centre-point replicates. Centre points are experiments where each factor is set to the midpoint value in its upper and lower bounds. These points serve as an assessment of the repeatability of the experiments performed to estimate experimental error within the system and can be further replicated to enhance this. An example FFD for a 2-level 3-factor design including a singular middle point, where the upper and lower bounds are set to +1 and -1 respectively, is shown in Figure 7, where *N* can be calculated to produce the requirement for 9 experiments.

$$N = n^k + m \qquad (4)$$



Experiment	X_1	X_2	X_3
1	-1	-1	-1
2	+1	-1	-1
3	-1	+1	-1
4	+1	+1	-1
5	-1	-1	+1
6	+1	-1	+1
7	-1	+1	+1
8	+1	+1	+1
9	0	0	0

Figure 7. Example 2-level (*-1, +1*) 3-factor FFD (*X1, X2, X3*) with a singular centre point. The constructed matrix for the experiments is highlighted on the right. Blue points represent experiments required from this matrix, with the centre point highlighted in red, originating from the data entry for experiment 9.

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_{12}X_1X_2 + b_{13}X_1X_3 + b_{23}X_2X_3 + b_{123}X_1X_2X_3 + \epsilon \quad \textbf{(5)}$$

Where Y is the response variable, X_i is an input variable, b_i is the associated contribution and ϵ is the related error in the model.

Often one experimental design is not enough, whereas in an optimisation process, an initial screening design might be performed to rule out any variables which have negligible effects. This would then be followed up with an optimisation design on the remaining variables. The data collected from the previous design can be used to regenerate and modify the problem to improve the next model. Modification of the postulated model, removal of insignificant factors, re-examination and definition of the experimental design can all be performed and repeated from point 3 (Planning of reactions) to enhance the new design being performed. This limits the overall process, where refinements of the optimal region require additional execution of experiments, thus increasing the experimental demand for process understanding to be achieved. Furthermore, for higher dimensional systems, DoE can suffer further from the necessity for an exponential number of required experiments to be performed. A combination of both these effects

can relate to significantly increased cost and time requirements for experimental procedures.

1.3.1.3 Simplex

Access to different types of coding software packages such as MATLAB and LabVIEW has helped give research chemists the ability to control laboratory equipment. Various types of algorithms have been implemented into this software which has enabled researchers to gain access to self-optimisation techniques and setups.

Nelder-Mead Simplex (NMSIM) is an example of a black-box optimisation technique where directed iterations to consistently improve an objective are performed and can be implemented on systems where DoE might become inefficient. A black-box technique can be described as an optimisation where no *a priori* knowledge or mechanistic understanding of the chemical system is required.[35] The algorithm can determine the maximum response of a single objective by using a numerical evaluation of the objective function.[35] The use of algorithms within optimisation can allow for a reduction in the number of experiments required to find the optimum conditions. A reduction in the number of experiments means that these processes are faster and more efficient, with reduced material consumption. The basis of the NMSIM optimisation algorithm stems from geometry by utilising a polyhedra that is formed of $n+1$ vertices, where n is the number of variables being observed.[36] This builds upon the original simplex algorithm which was first developed by Spendley *et al*. to include the following additional geometric transformations: multiple contraction of the polyhedra, outside contraction, reflection and expansion as shown in Figure 8.[37] These modifications were made to reduce the rigidity to overall improve the optimisation algorithm.[38]

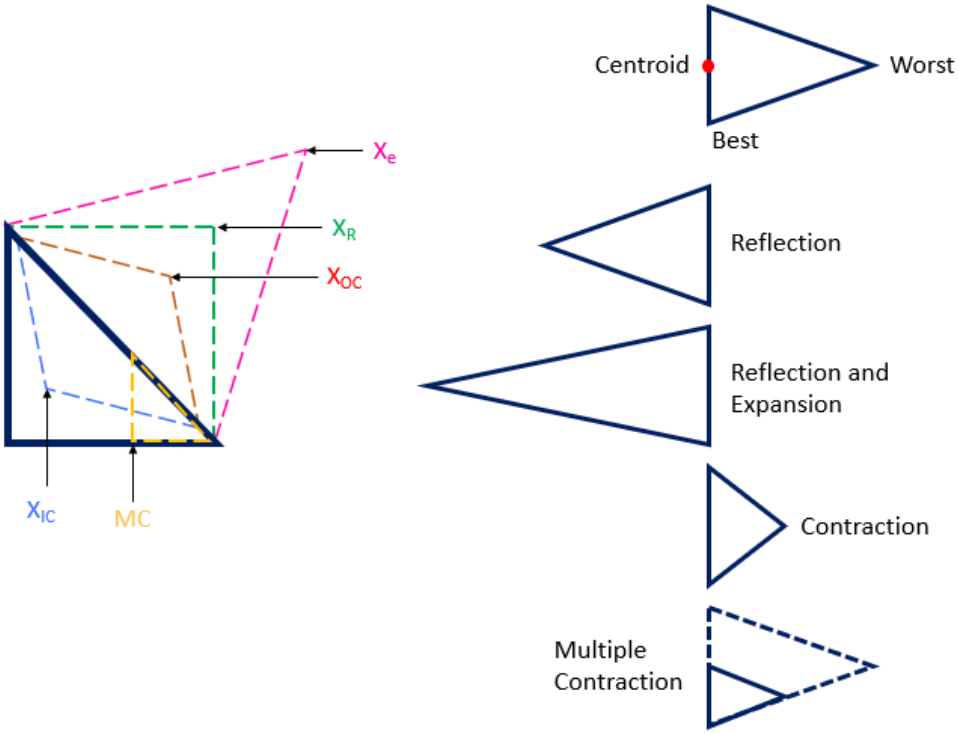
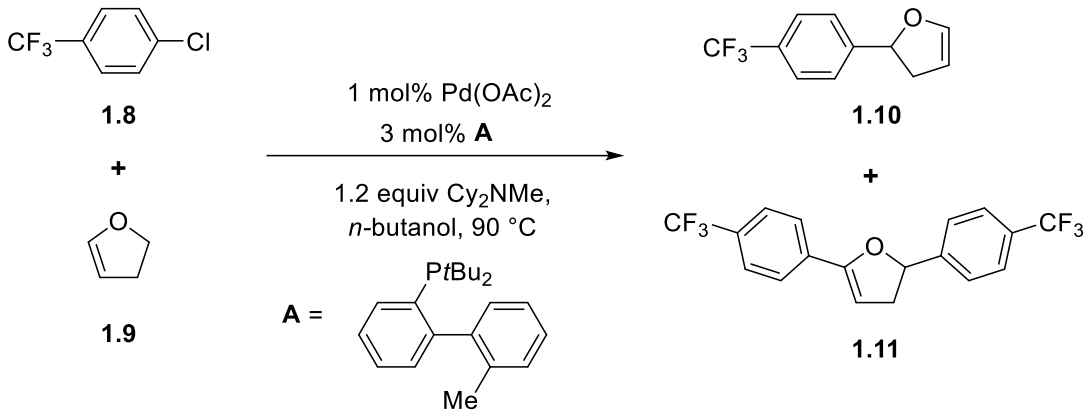


Figure 8. Possible geometric transformations performed by the Nelder-Mead Simplex algorithm for a 2D problem: multiple contraction (MC), inside contraction (X_{IC}), outside contraction (X_{OC}), reflection (X_R), expansion (X_e).

The algorithm explores the feasible design space set by either the user-defined variable boundaries or randomly generated experiments. Each vertex of the polyhedral generated represents an experiment where the response function of it has been evaluated. Upon each iteration, the vertex producing the worst response function is replaced with another vertex via a geometric transformation. This results in a completely new simplex that is exploring a new area of the design space. Repetition of this process continues with the objective to iteratively find vertices with better response values until it can find the local optima of the system.[39]

Jensen *et al.* employed the NMSIM algorithm for the optimisation of a Heck reaction between an aryl chloride (**1.8**) and 2,3-dihydrofuran (**1.9**) utilising a flow-microreactor system.[40] The case study aimed at optimising the Heck reaction's yield for the monosubstituted product (**1.10**) on the microscale by adjusting residence time and alkene (**1.9**) equivalents. The algorithm successfully identified optimal conditions to achieve an overall yield of 83% in just 19 experiments.



Scheme 2. Heck reaction between 4-chlorobenzotrifluoride (**1.8**) and 2,3-dihydrofuran (**1.9**) to form the mono-(**1.9**) and bi-substituted products used in the Nelder-Mead Simplex optimisation.

Whilst this optimisation highlights the efficacy of employing the NMSIM algorithm on chemical systems, the deliberate simplification of the process should be noted. Reducing the number of continuous variables examined constrains the number of variable-variable interactions within the system and as such reduces the process understanding gained from the optimisation. Additionally, the reduction in the number of continuous variables facilitated the decrease in the number of experiments required to achieve the optimised results. Furthermore, as a prerequisite to the optimisation, a solvent, palladium catalyst and phosphine ligand screen were all carried out. Notably, to credit this study, the reaction was successfully scaled up 50-fold using the optimum conditions achieved from the microscale process, with the conservation of comparable yields on the mesoscale.

Numerous local optimisation algorithms, including gradient-based methods and further modifications to the simplex algorithm, have been developed and implemented for the optimisation of chemical systems.[41–43] Individually detailing each category of local algorithms available for optimisations of chemical systems extends beyond the scope of this work. Table 1 provides an overview of self-optimisations that have been completed on chemical systems using local optimisation algorithms for the consideration of continuous variable case studies.

Table 1. Overview summary of examples utilising local optimisation algorithms on chemical systems.

Year, Group	Algorithm	Reaction	Variables	Objective
2010, Jensen[40]	NSIM	Pd-catalysed Heck reaction	Equivalents of alkene & Residence time	Yield
2010, Jensen[26]	NSIM, The steepest descent method, SNOBFIT	Knoevenagel condensation and benzyl alcohol oxidation	Temperature & Residence time	Objective function value & Associated yield
2011, Poliakoff[44]	Super Modified Simplex (SMSIM)	Methylation of 1-pentanol in $scCO_2$	CO_2 Flow rate, Temperature, Pressure, Equivalents	Yield
2011, Poliakoff[45]	SMSIM	dehydration of ethanol over γ-alumina in $scCO_2$, carboxymethylation reaction in $scCO_2$	Temperature, Pressure & CO_2 Flow rate	Yield
2012, Jensen[46]	The steepest descent method, Conjugate gradient, Armijo conjugate gradient and Penalised Armijo conjugate gradient	Paal−Knorr	Temperature & Residence time	$\frac{conversion}{residence\ time}$ (Proportional to Productivity)

2012, Poliakoff[47]	SMSIM	Methylation of alcohols using dimethyl carbonate (DMC) in $scCO_2$	Flow rates of CO_2, 1-pentanol and DMC Reactor temperature & Pressure	Yield STY E-factor E+ STY x Y
2013, Poliakoff[48]	SMSIM, SNOBFIT	Solvent-free methylation of alcohols with DMC over γ-alumina	Flow rates of 1-pentanol and DMC Reactor temperature	Yield
2015, Poliakoff[49]	SMSIM	γ-alumina catalytic reaction with aniline using DMC and THF in $scCO_2$	Analine, DMC and THF* flow rates, Temperature & Pressure * THF flow rate not optimised for first 4 products	Yield of a range of products within the reaction
2016, Felpin[50]	Modified Simplex	Pd catalysed Heck-Matsuda reaction	Catalyst loading (mol%), Temperature, Residence time & Equivalence of starting material	Yield, Highest throughput and production cost
2016, Ley[51]	Complex Simplex	Hydration of nitrile to amide over MnO_2, Appel reaction	Residence time, Temperature, Concentrations & Equivalents	Weighted objective function determined by: Throughput, Conversion, Consumption
2018, Felpin[52]	Modified Simplex	allylation, [3,3]-Claisen rearrangement, isomerisation,	Temperature, Residence time & Equivalents of starting material	Yield & Throughput

		oxidative dimerization		
2018, Ley[53]	Complex Simplex	(±)-tramadol synthesis (Grignard addition), lidocaine synthesis(acylation of aniline derivative, amine alkylation), bupropion synthesis (α-bromination, amine alkylation)	Residence time, Temperature & Equivalents	Four-term objective function determined by: Throughput, Conversion, Consumption, Energy
2018, Rueping[54]	Modified Simplex	Paternò-Büchi [2+2] photo-cycloaddition	Flow rates and residence time	Conversion
2021, Hein[55]	Bayesian Optimisation	Stereoselective Suzuki-Miyaura coupling	Temperature, Equivalents, Ligand : Catalyst ratio, Catalyst loading (mol%)	Four-term objective function: E-product assay yield (1st), Z-product assay yield (2nd), Pd loading (3rd), acid equivalents (4th)
2023, Bourne[20]	Bayesian Optimisation Adaptive Expected Improvement	Telescoped Heck-Selective Hydrolysis	Temperature, Equivalents, Residence time, Ratio of flow rate of acid to flow rate of first reactor	Overall product % yield

1.3.2 Global Optimisation

The methodologies presented in Table 1 are all examples of local optimisation techniques, where a local optimum in the design space is found. These types of optimisation techniques are typically fast as each iteration is converging on the optimum area. However, a disadvantage of local optimisations is when there are complex variable-variable interactions, this can lead to multiple local optima being observed within the system. The presence of multiple local optima in a system can cause a local optimisation algorithm to get stuck on a single optima and home in towards it, when in fact it might not be the true optimum value, Figure 9.

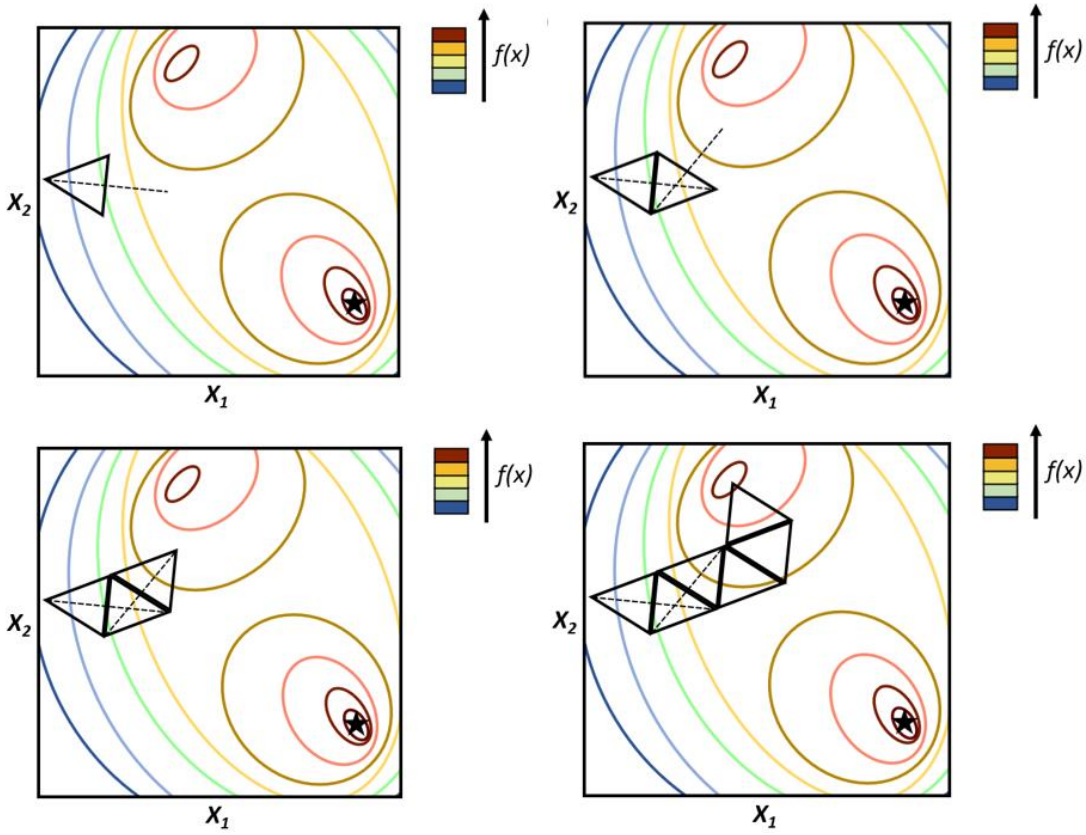


Figure 9. A visual example of a 2-variable simplex optimisation converging on a local optimum where a global optimum of a better solution is present.

From these challenges, when attempting to employ a local optimisation algorithm into a problem, it must be assured that the specific chemical system only contains one local optimum, which can be obtained by limiting the number of variables under study. Global optimisation algorithms possess the ability to effectively manage noise in complex systems, allowing them to achieve faster

and more reliable convergence towards optimal solutions compared to local optimisation techniques like Simplex, especially in noisy systems.[56]

1.3.2.1 Stable Noisy Optimisation by Branch and Fit

The Stable Noisy Optimisation by Branch and Fit (SNOBFIT) method is another example of a global optimisation algorithm which was developed by Huyer and Neumaier.[56] Once again, it is an example of a black-box technique as well as being a derivative-free method, meaning that it does not require any gradient information on the objective that is being optimised. The algorithm uses a combination of stochastic linear and quadratic surrogate models to find the global optimum of the system.[35] These surrogate models allow for the prediction of the black-box model and predict the outcome of interest where a limited amount of data is available, which can provide a cheaper alternative to conducting reactions.[57]

Within the optimisation, the algorithm generates 5 different classifications of points which are detailed below. A flow diagram is detailed in Figure 10 where each classification generated is included.

Class 1 - The point that minimises the local quadratic around the best current point and contains at most one point.

Class 2 - Points that approximate local minimisers. Where no local points are generated within an optimisation, no points in class 2 are generated.

Class 3 - Points that are approximate nonlocal minimisers.

Class 4 - Points in unexplored regions.

Class 5 - Randomly generated points to fill the design space.

The main advantage of the SNOBFIT algorithm is that it can handle noise effectively due to the random fitting of points which leads to higher confidence that the optimum found within the optimisation will be the global optimum. However, due to the algorithm exploring the experimental area, this can cause a higher number of experiments to be required such that it is outperformed in terms of number of reactions required by local optimisation techniques for simple (no noise) systems.

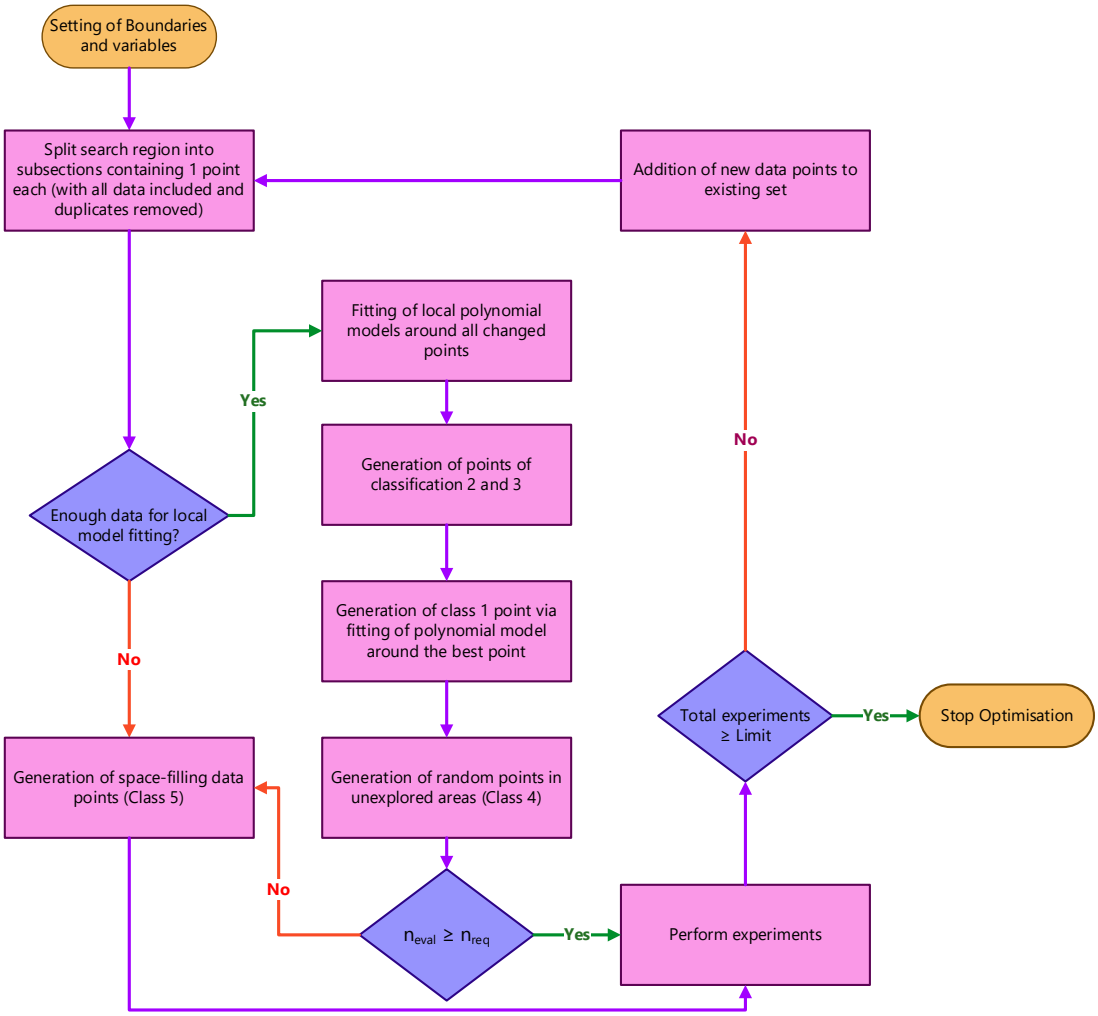


Figure 10. Flow diagram for a SNOBFIT optimisation. Where $\boldsymbol{n_{req}}$ = minimum number of required function evaluations to be performed, set by the user, $\boldsymbol{n_{eval}}$ = number of actual function evaluations performed.

SNOBFIT has been implemented for the optimisation of chemical systems on several occasions across various groups, with the Bourne group significantly contributing to this literature.[49–51] The first example of SNOBFIT being implemented was by Krishnadasan *et al.* for the controlled synthesis of nanoparticles.[58] Optimisation took place by measurements of the wavelength of the outlet of the reactor until the desired nanoparticle size was formed. The injection rates of reactants and the temperature of the reactor were controlled by the algorithm to perform a noise-tolerant global search method.

However, SNOBFIT does begin to struggle with chemical systems where the number of dimensions involved in the optimisation is too large (>9 dimensions).[61] Telescoped reaction optimisations pose an obvious problem where multiple reagent streams, reactors and therefore dimensions are

involved. As well as high dimensionality problems, the algorithm can only be applied to optimise continuous variables, so for the optimisation of ligands, bases catalyst etc. a different type of optimisation algorithm will have to be employed.

1.3.2.2 Bayesian Optimisation

Bayesian optimisation represents a division of derivative-free global optimisation methods that employ probabilistic models for an initially unknown objective function. It is designed for optimising functions that are expensive to evaluate.[62] Bayesian optimisation offers advantages by enabling the incorporation of prior knowledge about the problem, guiding the sampling process to explore a balance between exploration and exploitation of the design space. These Bayesian optimisation methods rely on surrogate models to approximate the true objective function. In each iteration, the next point for evaluation is determined by an acquisition function, which is calculated using these constructed surrogate models generated from prior obtained data. Typically, the most common surrogate models take the form of Gaussian processes (GP), which define a distribution over all possible functions in line with the observed data, [Eq (**8**)(**6**)].[63] Specifically, a GP can be fully defined by a mean function, $m(\mathbf{x})$, [Eq (**6**)], and a covariance function, $k(\mathbf{x},\mathbf{x}')$, [Eq (**7**)], of a real process $f(\mathbf{x})$, where **x** and **x'** are input vectors and $\mathbb{E}[\cdot]$ is the expectation over the function.

$$\mathrm{m}(\mathbf{x}) = \mathbb{E}[\mathrm{f}(\mathbf{x})] \quad (\mathbf{6})$$

$$\mathrm{k}(\mathbf{x},\mathbf{x}') = \mathbb{E}[(\mathrm{f}(\mathbf{x}) - \mathrm{m}(\mathbf{x}))(\mathrm{f}(\mathbf{x}') - \mathrm{m}(\mathbf{x}')) \quad (\mathbf{7})$$

$$\mathrm{f}(\mathrm{x}) \sim \mathrm{GP}(\mathrm{m}(\mathbf{x}), \mathrm{k}(\mathbf{x},\mathbf{x}')) \quad (\mathbf{8})$$

The covariance function, $k(\mathbf{x},\mathbf{x}')$, often referred to as the kernel, is used to determine the relationship between pairs of inputs to establish confidence intervals for the mean function. The kernel contains various hyperparameter settings, one of which is an adjustable setting that can be fine-tuned to accommodate noisy optimisations, where optimal values can be generated through the training of the model.[63,64] Upon acquisition of data, the prior distribution evolves into a posterior distribution with updated beliefs about the unknown system. This posterior distribution is utilised to formulate the

previously mentioned acquisition function. The acquisition function is used to determine the subsequent evaluation point based on the trade-off between exploration and exploitation.[65] Experiments are suggested by the acquisition function to reduce uncertainty in the most promising regions. The acquisition function attains higher values when: (i) Mean function/exploitation, is high: (ii) Uncertainty/exploration is high: (iii) or both points are true: in cases involving a maximisation optimisation.

1.3.2.3 Multi-Objective Optimisation

The concept of multi-objective optimisation describes the methodology aimed at optimising a process by considering various performance metrics simultaneously. In chemical systems, a wide variety of criteria exist for evaluating the performance of a reaction, including environmental factors like the E-factor or reaction mass efficiency (RME), as well as productivity-related metrics such as Space-time yield (STY) or yield.

In contrast to single-objective optimisation, which focuses solely on one performance criterion, often favouring productivity metrics, it is evident that in many chemical systems, optimising the throughput of a system does not always translate to improved reaction performance. This, in turn, can lead to substantial waste generation.[66] These approaches have been employed in a sequential approach where individual process metrics are consecutively optimised independently of one another to identify optimum conditions for each objective. Consequently, these methodologies lack consideration of how the objectives interact and the essential interactions of the variables required to explore a trade-off between them. An alternative strategy involves the combination of multiple objectives into a single-weighted objective function, allowing users to assign weights and preferences to process metrics, which are then optimised. However, such workflows often require *a priori* knowledge of the chemical system’s various metrics, posing challenges in practical implementation where slight adjustments to these weighting values can lead to drastic output variations. Furthermore, weighted objective optimisations only yield a single optimal solution, failing to explore the trade-off between objectives.[35]

Conversely, the alternative is an *a posteriori* approach which involves the simultaneous optimisation of a combination of multiple conflicting objectives, intending to achieve a set of non-dominated solutions that explore the trade-off between objectives. These types of methodologies do not rely on assigned weights or bias towards different objectives unlike its a *priori* counterpart. Simultaneously optimising these objectives within a single process provides insight into the unique interactions between variables that are required to optimise each metric and explores the inherent trade-offs between them.[50] In these systems, due to the intrinsic conflict between these objectives, single optimal solutions are often not generated. Instead, a Pareto front, defined as a set of optimal non-dominated solutions that elucidate the trade-off between the objectives, is generated, Figure 11. [35,67]

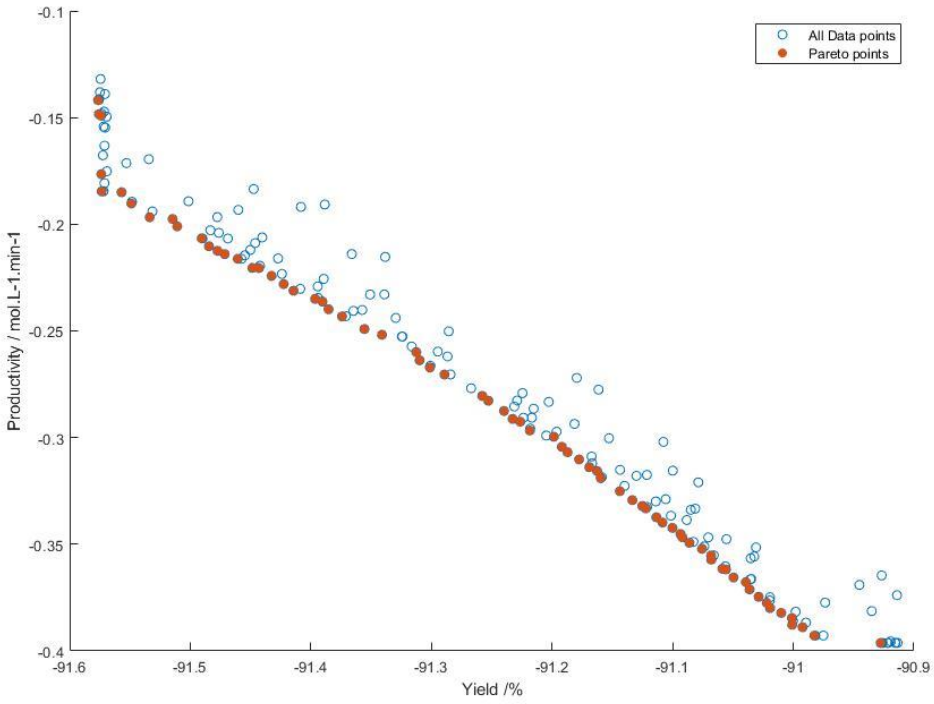


Figure 11. An example of a conflicting objective optimisation, between productivity/STY and Yield, where both metrics are being minimised. Red points illustrate the Pareto front between the objectives, with blue points being dominant solutions.

There are several common criteria used to evaluate the performance of a multi-objective optimisation process, with the hypervolume of a system being the most commonly used. This metric serves as a measure of how

effectively a collection of solutions spans the Pareto front, where a greater value of hypervolume implies superior coverage suggesting a more diverse and desirable set of solutions.[68,69] The interest and popularity of hypervolume as a main performance criterion stems from the fact it considers both the precision and diversity of an approximate set. These associated beneficial attributes of the hypervolume for measuring accuracy and diversity of the dataset, led to the utilisation of this metric to determine optimisation termination that is featured in Chapters 2-4.

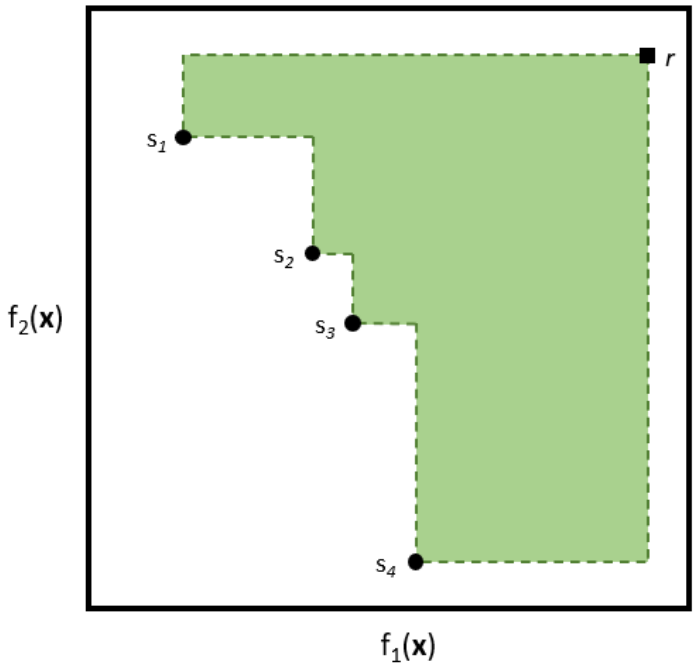
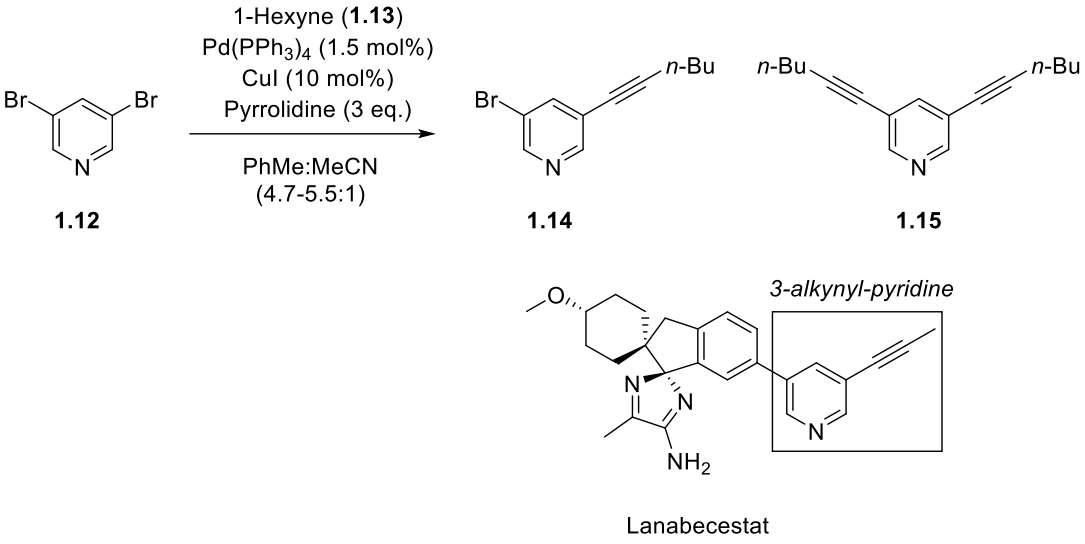


Figure 12. Hypervolume plot for a 2D objective space. The hypervolume is calculated between the reference point, *r*, and the current non-dominated solutions, s_i.

Multi-objective optimisation approaches have been well established in chemical systems as evidenced by a range of published literature examples.[66,70,71] Clayton *et al.* highlighted the application of a Thompson sampling efficient multi-objective optimisation algorithm (TSEMO) on a Sonogashira reaction for the synthesis of the 3-alknyl-pyridine moiety in lanabecestat, which is a drug used for Alzheimer's disease, Scheme 3.[72]



Scheme 3. Sonogashira coupling reaction optimised using a TSEMO algorithm between 3,5-dibromopyridine (**1.12**) and 1-hexyne (**1.13**) to form the desired mono-substituted alkyne (**1.14**) and undesired bis-alkyne (**1.15**).

The Bayesian algorithm utilises an initial dataset, generated via a random set of space-filling experiments, to construct GP surrogate models for the optimisation process. A random sampling of the GPs is utilised by the algorithm to ensure a balance of exploration and exploitation is achieved within the optimisation and further reduces bias towards any objective under study. The sample with the highest predicted hypervolume improvement is selected for the next sample and is subsequently performed in Figure 13.[35]

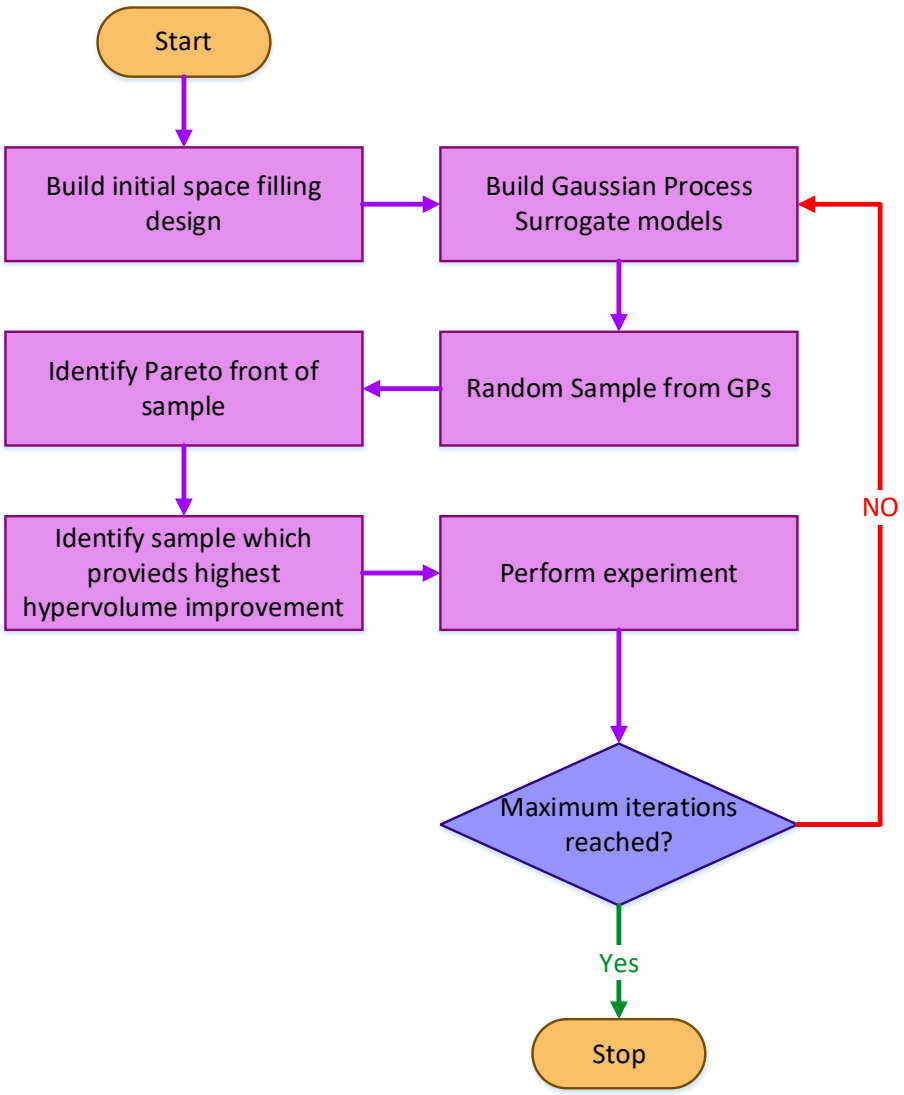


Figure 13. Flow chart for the TSEMO algorithm workflow within an optimisation process, utilising random sampling of GP surrogate models.

1.4 Mixed Variable Systems

While numerous studies have delved into optimising continuous variable systems to maximise process metrics, the exploration of mixed variable systems remains relatively unexplored. In recent years, optimisations have increasingly focused on integrating both discrete and continuous variables, enabling a comprehensive examination of variable interactions within chemical systems.[73–75] However, many industries and synthetic research still employ the traditional OVAT-based methodology for the screening of discrete variables within their optimisation processes. The vast combination of potential discrete variables within a chemical system can lead these workflows

to be highly resource-intensive. Furthermore, the inclusion of discrete variables within an optimisation process naturally increases the complexity of the system due to the extended networks of variable interactions possible, resulting in an increased likelihood that the true optima of the system is not identified. High-throughput experimentation (HTE) has been well-documented to perform large numbers of automated experiments in the pharmaceutical industry.[76–78] HTE can be achieved using auto-sampling platforms that can perform multiple parallel reactions simultaneously. Using this, a wide library of compounds can be screened to identify the optimum discrete variable (catalyst, ligand etc.) at the current operating conditions.[79] The method is especially useful for the minimisation of waste produced where reactions can be screened on a small scale to ensure minimal amounts of material are required.

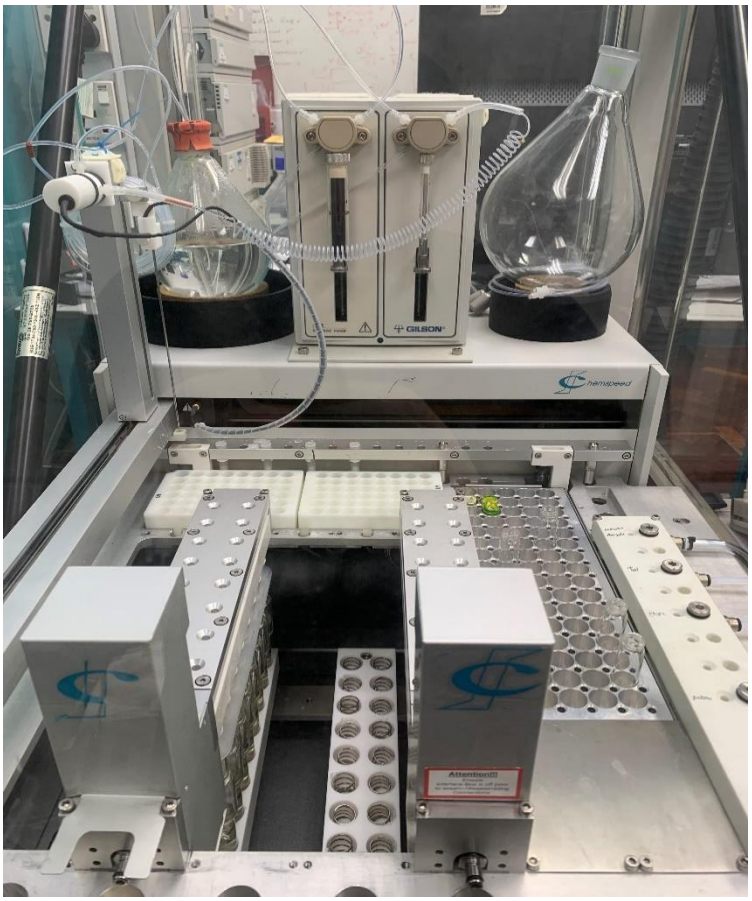


Figure 14. Photo of a Chemspeed HTE batch reactor, where aliquots of mixtures are drawn up using syringe pumps and injected into the batch vessels.

The combination of HTE, coming through an autosampler, combined with continuous flow provides a very powerful technique for the automated screening of compound libraries. Perera *et al*. highlighted the efficacy of this methodology on a Suzuki-Miyaura coupling reaction, in which the screening of suitable electrophiles, nucleophiles, bases, solvents and ligands was achieved, Figure 15.[73] In this method, 5 µL slugs of the desired discrete variable were injected into 500 µL of carrier solvent such that a 1:100 diffusion ratio was achieved to ensure the sufficient dilution of the reaction slug into the carrier solvent for the further downstream reaction.

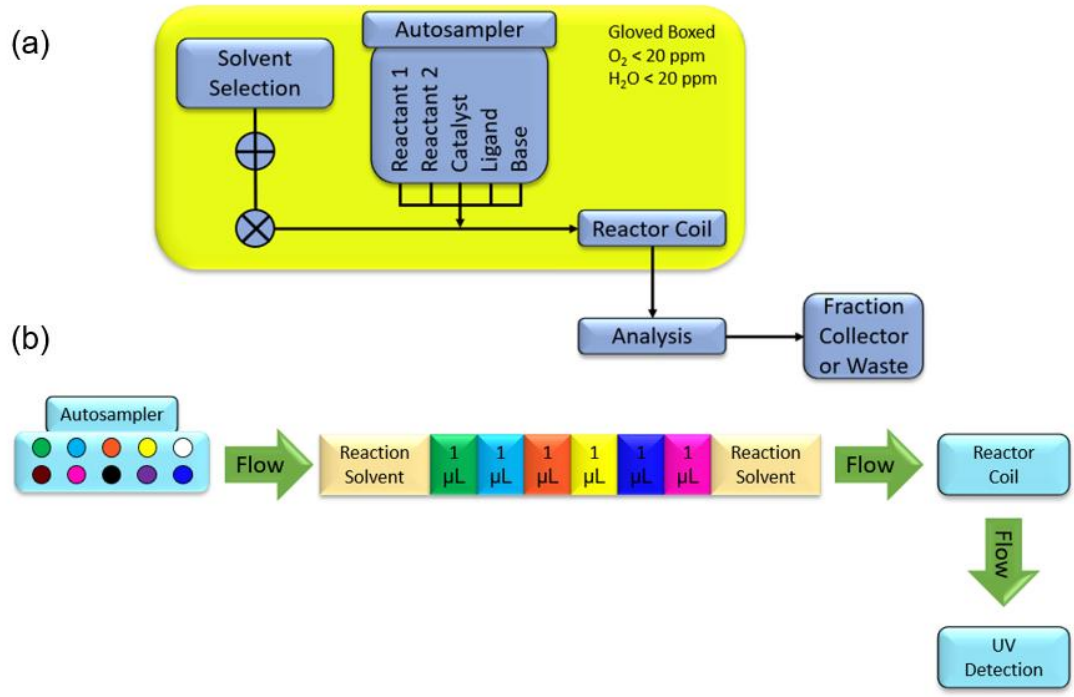


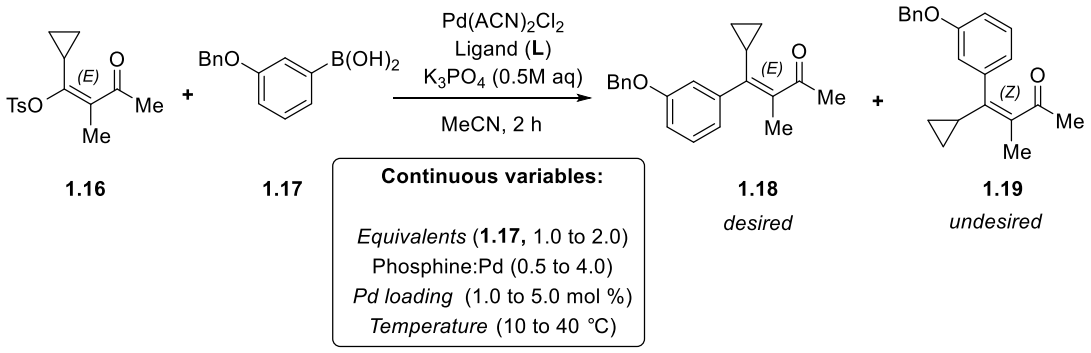
Figure 15. (a) Schematic of the autosampler-flow platform used to screen libraries of discrete variables for the Suzuki-Miyaura cross-coupling reaction. (b) Preparation and injection of discrete variable slugs into the flow stream, where mixing and diffusion it analysed by UV detection.

Combination with continuous flow facilitated the precise control of reaction parameters, where the flow rates, residence time, reactor temperature and system pressure were all regulated. The inclusion of the autosampler in the system enabled a 1 µL aliquot for each discrete variable (nucleophile, electrophile, catalyst, ligand and base) to be injected to form the 5 µL slugs in the carrier solvent in a fully autonomous manner. Additionally, utilising an autosampler for this system facilitated the flexibility of placing it inside a glovebox such that it provided access to moisture or air-sensitive

discrete variables. Using this platform, a total of 5760 experiments were successfully performed achieving a rate of >1500 reactions in 24 hours.

Although this flow-HTE platform was highly effective for screening a large library of discrete variable compounds, the process eliminated the consideration of continuous variable interactions within the system via the selection of a single set of operating conditions. Furthermore, this workflow presented a brute-force approach, where no optimisation was performed, instead a vast combination of discrete variables was tested over the process and as such required a significant number of experiments. Consequently, this limitation is highly problematic due to the complex interactions between variables, which have been shown to significantly affect reaction outcomes. Therefore, this leads to a desire for the simultaneous consideration of mixed-variable optimisations on chemical systems, especially for cases with the consideration of multiple objectives.

There has been a singular investigation for the consideration of mixed-variable optimisations concerning multiple objectives performed by Christensen *et al.* employing a Chemspeed equipped with an online HPLC to autonomously conduct and analyse batch-based experiments.[55] In this workflow, a stereoselective Suzuki-Miyaura coupling was selected as the chemical system, where various phosphine ligands were utilised as the discrete variable of choice, Scheme 4.



1.16 **1.17** **1.18** *desired* **1.19** *undesired*

Scheme 4. Suzuki-Miyaura cross-coupling reaction between *(E)*-vinyl sulfonate (**1.16**) and aryl boronic acid (**1.17**) used for the mixed-variable optimisation where maximisation of the *(E)*-desired product (**1.18**) was priority.

This work presented the first case of a mixed variable multi-objective optimisation on a chemical system, that was directed by a machine learning

algorithm. However, a significant disadvantage of this method was its reliance on batch-based experimentation, which imposed a maximum temperature limit of 40 °C, thereby mitigating any potential interactions between variables and elevated temperature ranges. Furthermore, using batch processing for the optimisation meant that each experiment was conducted for a set 2-hour period, leaving unexplored possibilities for varying reaction times. Given the inherent complexity of mixed variable systems and the restrictions imposed by the limiting temperature range and reaction times, this approach limited the process understanding gained from the optimisation. In contrast, adopting a continuous flow system would address these limitations by enabling access to higher temperatures and various residence times, thus facilitating a comprehensive exploration of interactions across a wider range of conditions. The exploration of mixed variable optimisations within continuous flow processes has been limited, with previous evaluations focusing solely on single-objective systems. Further inclusion of details regarding the optimisation of discrete variables within flow processes is discussed in Chapters 2-4 introductions.

1.5 Discussion

The exploration of discrete variables in research is not a new frontier. Many organic synthetic procedures have historically employed OVAT methodologies to identify the “optimal” operating conditions and build a library of functional groups for these chemical transformations. However, the OVAT methodology has proven unreliable in identifying the true optima of chemical systems where complex variable interactions are present and are often labour-intensive, demanding a large number of reactions. The use of HTE has also been well-published for the autonomous screening of vast libraries of compounds for the synthesis of APIs within the pharmaceutical industry. Nevertheless, there is a trade-off, as the time required per reaction can be increased, when compared to flow reactions, due to the temperature constraints, typically related to solvents, for batch processes.[76,78,80,81] However, the adoption of flow systems for these processes provides access to higher temperatures and enables the exploration of hazardous

intermediates within these reaction pathways to further enhance early-stage reaction development.

Over the past decade, numerous literature examples have featured self-optimisation platforms for a variety of chemical systems to determine optimal conditions concerning a range of process metrics. These studies typically examine the exploration and utilisation of different algorithms to control experiment procedures and streamline the optimisation process. Local optimisation techniques, such as Simplex and gradient-based methods have been significantly exploited for chemical systems with low-level complexity.[26,44,45,82,83] Similarly, global optimisation methods, such as SNOBFIT, have demonstrated successful approaches in exploring more complex chemical systems.[58–60,66,84–86] However, all of these previous approaches have only examined the optimisation of continuous variables within these flow systems, mitigating any potential exploration of interactions between discrete and continuous parameters. The Jensen group has provided the main contribution to literature for consideration of mixed-variable flow systems, where only single-objective optimisations were explored.[74,79,87]

1.6 Project Aims

The combination of automated flow platforms with efficient optimisation algorithms has facilitated the expansion of the toolkit available for process development. Furthermore, the utilisation of Bayesian algorithms has proven successful in expanding the capabilities of these processes for the consideration of multi-objective optimisations. However, there is an absence of the simultaneous optimisation of mixed variable flow systems concerning multiple objectives. There is only a singular case documented examining mixed variable multi-objective optimisations, with the workflow relying on batch processing to achieve this, where the previous limitations of this system related to temperature have been highlighted.[55] Therefore, presenting a necessity for these workflows to be developed into automated continuous flow systems, to provide an efficient approach to comprehensive reaction optimisation. Implementation of these optimisations will provide a crucial

expansion to the tools available for process development, especially in early-stage reaction development.

Chapter 2 initially establishes the potential for simulated optimisations of chemical systems using pre-existing kinetic data to provide insight into an S_NAr case study that would subsequently be examined for experimental optimisations. Building on these simulations, work was completed to highlight the experimental applicability of mixed-variable multi-objective optimisations on chemical systems using automated continuous flow processing. Two conflicting objections were set such to explore the solvent effects on the regioselective outcome of the S_NAr, where the solvent of the system was selected as the discrete variable enabling full exploration of the trade-off between the process metrics.

Chapter 3 builds on the previous mixed-variable optimisation to present the potential of these workflows for early-stage reaction development in the pharmaceutical industry, such that the synthesis of an API was examined. The inclusion of a Sonogashira cross-coupling reaction enabled a more chemically complex catalytic system to be optimised to further explore more complex variable interactions within the system and enhance the applicability of these methodologies. The optimisation was set to explore the trade-off between an environmental and productivity metric such that the ever-increasing concern about the environmental effects stemming from the production of APIs would be considered.

Chapter 4 describes the further development of the automated flow platform to introduce the optimisation of mixed-variable telescoped reactions. Multi-point sampling was implemented to facilitate the analysis of all steps within the telescoped pathway on a single piece of analytical equipment. Conflicting objectives were selected from each reaction within the system to enable different reaction pathways to be optimised, where the inclusion of a Heck-hydrolysis reaction further highlighted the development of catalytic systems. A selection of mono- and bidentate phosphine ligands was included as the discrete variable to examine the potential mechanism of the Heck reaction and the outcome effects on the conflicting process objectives.

Chapter 2. Synergising Simulation and Experimental Studies for Mixed Variable Optimisation

2.1 Introduction

The COVID-19 pandemic significantly highlighted the importance of being able to conduct simulated experimentation, specifically in the optimisation landscape which can excel in efficiency and effectivity for these types of processes.

The pandemic had a significant impact effect worldwide on the ability of chemical researchers to conduct experiments, leading to many research groups finding alternate ways to carry out work. Still being able to carry out 'experiments' would be vital to gaining insight into potential future case studies, further developing PhD project timelines, and contributing work to chemical literature. Computation chemistry is already a well-known and researched branch of chemistry, that utilises mathematical methods in combination with chemical physiological and electrical properties to build models and approximations that can be implemented on a computer.[88] Simulated optimisations combine the power of computational software packages with efficient mathematical algorithms to explore the chemical system scope and predict the outcomes to enable researchers to have critical insight into how the chemical system under study will behave.[89]

Developments in kinetic analysis software, such as Compunetics, over the recent years have enabled this area of chemistry to become readily available to the broader research community.[90,91] The software can apply the mathematical principles and models to the chemical systems using the inputs of real-world kinetic data that have been previously gathered.[92–94] With a vast amount of kinetic data readily available through published literature work, this allows for simulated computational experimentation to be carried out. This data allows for real-world kinetic models to be generated from all known species in the specified chemical reaction under study, to investigate how the properties of each substituent in a reaction affect the outcome. Performing these types of simulated experiments can provide key insights into quantitative chemical synthesis for process development which poses a critical constraint for the chemical industry, that when utilised can reduce the time requirements and costs for the scale-up of these processes.[95,96] MATLAB offers an inbuilt ordinary differential equation (ODE) solver into their software

which enables the models of the concentration profiles for each component in reaction to be simulated. These types of studies provide key insights into the mechanisms of the chemical processes and enable the exploration of the optimisation surface responses which can aid in limiting the number of experimental reactions required. By doing this, it enhances the efficiency and sustainability of these processes, via the reduction of the number of experiments required which decreases the time and material consumption.[97] With this additional time, researchers can then focus their attention on handling more challenging problems which might be limited by experimental conditions, whereby simulated experiments can mediate the handling of hazardous intermediates, dangerous conditions, or laborious exploration of the design space.

The mathematical optimisation algorithms can then be employed in the simulated models via the manipulation of the variables under study to optimise the desired process metrics, whether economical or environmental.

Table 2. Comparison of Environmental and Economic metric objectives for evaluating chemical processes.

Economical	*Environmental*
Yield	E-Factor
Productivity (Space-Time Yield)	Reaction Mass Efficiency
Cost	Atom Economy
Purity	Impurities

When combined with simulated experiments, real-world experimental optimisations can provide a validation of the previously performed simulated work, all whilst building upon the simulated models to further refine and enhance the models being used. Leading to an additive effect where both techniques benefit from one another in a synergistic form in which, simulated models are being improved upon by experimental reactions and experimental design spaces gaining key insights into the chemical system landscape through the simulated experiments. This synergy grants researchers an

extremely powerful optimisation tool to accelerate optimisation processes and access larger data sets to analyse and report on.

Over the last decade, the majority of optimisations on chemical systems have only studied the manipulation of continuous variables for their desired outcomes, with only mixed variable case studies being observed in batch systems through the use of HTE in a ‘brute force’ manner.[73,76,80,98] These types of operations come with significant problems as when confronted with ever more intricate synthetic methodologies, the time-consuming ‘exhaustive’ strategy for high-throughput screening (HTS), involving the trial of every conceivable combination of discrete factors, is losing its appeal. Moreover, the process of optimising both discrete and continuous variables sequentially within a workflow leads to an inadequate understanding of the overall process, Figure 16. This deficiency arises from the neglect of crucial interactions that can occur among the discrete and continuous variables. In addition, the optimisation of mixed variable systems for multiple objectives is an area of research that is significantly lacking, this could be due to the nature of the complexity of the systems or algorithms able to adequately handle these types of problems.

However, there has been work reported on performing simultaneous optimisation of both categorical and continuous variables using an HTS batch system.[55] These types of experimentations come with the drawbacks associated with performing experiments in batch reactors compared to continuous flow processes which offer access to higher temperature reactions and enhanced control over hazardous intermediates. Warren *et al.* investigated a multi-objective optimisation of this type for a series of RAFT polymerisations to explore molar mass dispersity and monomer conversion.[99] To achieve this, they employed a TSEMO-based algorithm in their system. However, it is important to note that separate optimisations were completed for each RAFT agent, consecutively, leading to a significant concentration of data in suboptimal areas within the reaction scope.

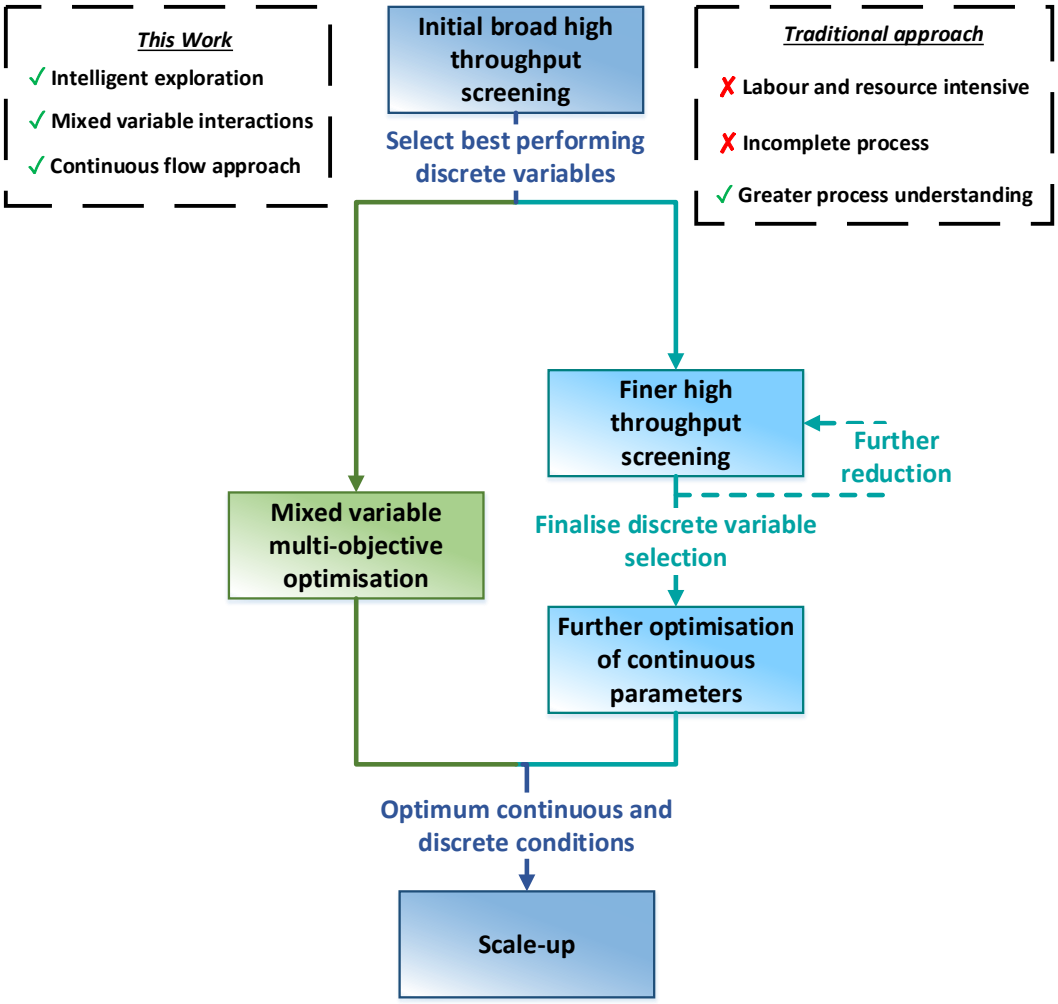


Figure 16. Comparison between traditional HTS methodologies vs. this developed MVMOO self-optimisation approach for a full reaction optimisation pathway.

Interactions between continuous variables in a chemical reaction pathway are becoming increasingly appreciated by synthetic research groups, with many procedures adopting new condition optimisation methods by combining computation and modelling techniques to revolutionize their research. The increase in appreciation could be due to the relative increase in literature publications of continuous variable studies and the ease of being able to define and optimise continuous variables compared to discrete variables.[6,47,52,86,100,101] Continuous variables are any variable within a reaction pathway that can be fitted to a continuous scale, and as such are typically easier to optimise with a wide range of optimisation techniques becoming readily available.

Discrete variables are typically harder and fit to a categorical scale, however, the complexity increases with discrete variables due to the number of properties (ligand steric bulk, catalysts ligand exchange rate etc.) that each variable can possess. The properties of each discrete variable can cause further variable-variable interactions to be observed within an optimisation space. These defined properties for discrete variables has prompted research aimed at categorising these factors with the goal of aligning them on a continuous scale, e.g. fitting solvents based on polarity index or ligands based on steric bulk. Jensen *et al.* used molecular descriptors to explore solvents to convert these discrete variables onto a continuous scale in a multi-objective optimisation of Rhodium-catalysed asymmetric hydrogenation in which the temperature was also optimised.[75] However, these methods require the exploration of extensive libraries of properties associated with the discrete variables under study. Selecting suitable molecular descriptors that are capable of accurately explaining the observed behaviour demands a substantial amount of prior knowledge and readily accessible data resources. This becomes particularly challenging when dealing with novel reaction pathways, as acquiring such knowledge can prove difficult to obtain without extensive reaction modelling and investigation. In contrast, black-box optimisation techniques circumvent the necessity for explicit physical knowledge of the discrete variables in the optimisation, bypassing this requirement of vast prior knowledge.[102]

Herein, this chapter aims to explore the simultaneous optimisation of mixed-variable flow systems, utilising the MVMOO algorithm to direct this process. The S_NAr reaction was selected as the initial case study, with a specific focus on optimising the yields of the various regioisomers formed.

The S_NAr reaction was selected as the initial case study to where the yields of the different regioisomers formed would be set as the objectives for the optimisation. This difference in regioselectivity of the reaction would be further tested via the selection of the solvent choice as the discrete variable, where previous literature has highlighted the different regioselective outcomes based on the solvent of choice.

Preliminary efforts will focus on leveraging simulated reaction chemistry and state-of-the-art machine learning algorithms to conduct simulated optimisations. This approach aims to offer insights into the S_NAr reaction, serving as supporting evidence for the experimental studies to follow.

2.2 Mixed Variable Multi-Objective Optimisation Algorithm

The mixed variable multi-objective optimisation algorithm (MVMOO) utilised within all experimental work contained within this thesis was created and developed by Jamie Manson as part of his PhD project within the group and was subsequently published in the Global Optimisation Journal.[63] The work within this thesis focuses on the implementation of the MVMOO algorithm onto real chemical systems, when used in combination with the automated flow platforms developed to optimise for multiple objectives during each case study. A brief analysis of the optimisation algorithms functionality is shown in Figure 17.

From Figure 17, the initial space-filling design was completed for each discrete variable, which was a five identical Latin hypercube (LHC) sampling for each categorical variable under study with varying continuous variable conditions. Using this methodology provides sufficient exploratory information in the design space for the algorithm for use with the iterative process models.[103] This type of procedure is suitable when working with small numbers of discrete variables but can lead to large initial costs to the optimisation when $n > 10$, where n is the number of discrete variables or when there are multiple different types of discrete variables being evaluated. Following this, individual GPs are utilized as surrogates for each objective function using the data set obtained from the initial conditions or current data set after each successful iteration. To enable the modelling of mixed variables, the GP surrogates use an internal distance metric which has been based on Gower similarity. During the iterative process, the next set of reaction conditions is determined via internal optimisation of the expected improvement matrix (EIM) acquisition function. An EIM function is used to train the GP models to determine the next point for evaluation.[104] The experiment is then carried out on the automated platform and the analysis is evaluated for

the objective functions. After performing the experiment, the GPs are updated after which the process proceeds in an iterative process until a stopping point is achieved. This can be determined by evaluating the hypervolume changes after a set number of iterations. Using this iterative approach enables a balance of exploration and exploitation to identify the global Pareto front for all metric functions.

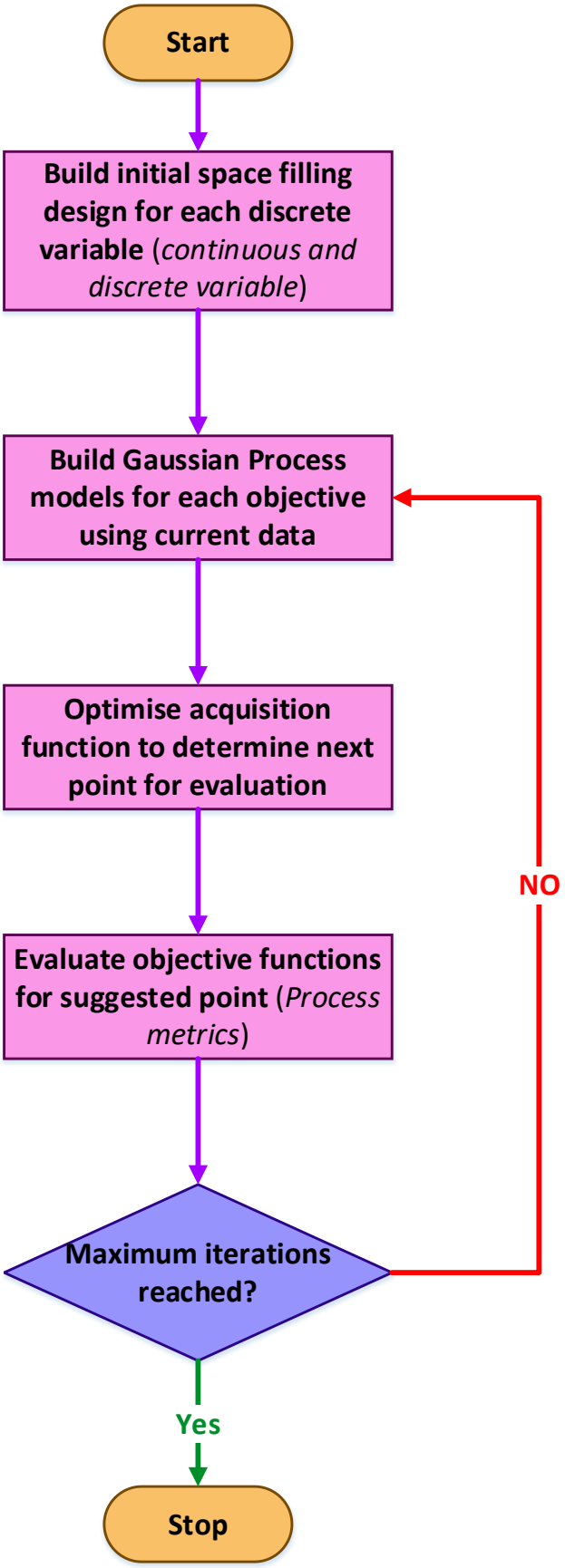
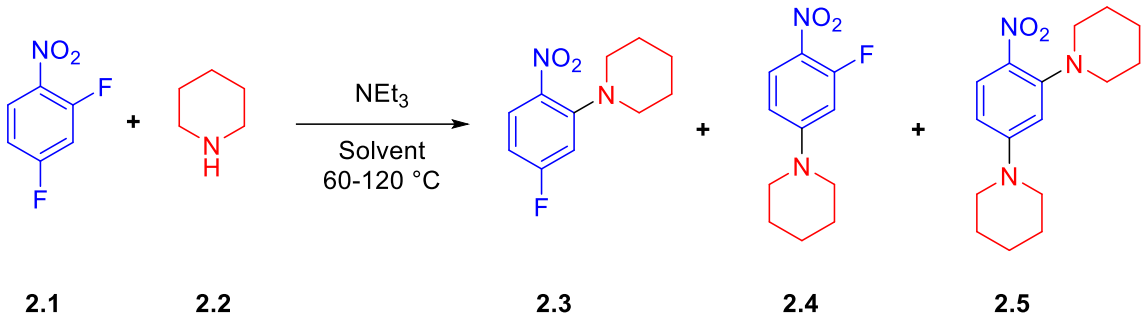


Figure 17. Flow chart for the iterative process of the MVMOO algorithm.

2.3 Nucleophilic Aromatic Substitution

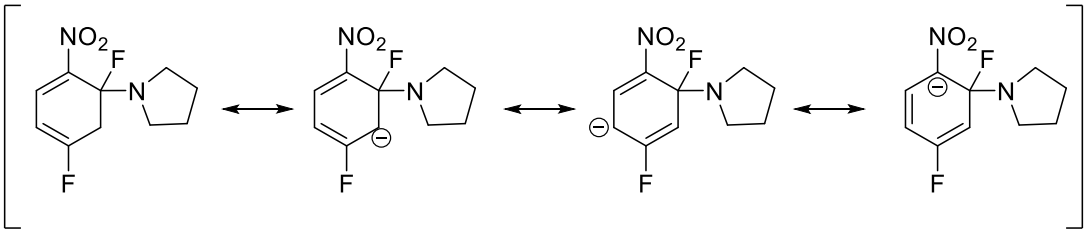
The nucleophilic aromatic substitution (S_NAr) reaction has been utilised as an ideal test reaction for continuous optimisation reactions of the chemical system for both single- and multi-objective studies.[66,105,106] For this optimisation, in the initial studies, operating conditions were utilised from the work completed by Schweidtmann *et al*. but were later subjected to some adjustments for improved stability of the system.[66]



2.1 **2.2** **2.3** **2.4** **2.5**

Scheme 5. S_NAr reaction between 2,4-difluoronitrobenzene (**2.1**) and morpholine (**2.2**) for the formation of the ortho- (**2.3**), para- (**2.4**) and bis- (**2.5**) regioisomers.

The S_NAr reaction serves as a valuable synthetic tool for substituting effective leaving groups on an aromatic ring, commonly using aryl halides. This process involves the utilisation of an appropriate nucleophile and unfolds through a resonance-stabilised Meisenheimer complex, illustrated in Scheme 6. This reaction proves advantageous for crafting aromatic carbon-heteroatom bonds, with the potential to create diverse regioisomers hinging on the selection of a suitable starting material. The ability to yield both desired regioisomeric products and undesirable products further designates this reaction as a key reaction for the initial case study. The inclusion of this reaction would help increase the complexity of the optimisation at hand to facilitate the assessment and refinement of algorithmic capabilities on the automated flow platform.



Scheme 6. Resonance stabilisation of the *ortho* regioisomer through a Meisenheimer complex.

When choosing an initial case study for the optimisation of mixed variables, it was logical to select the S_NAr reaction. This choice stemmed from the reaction's historical performance within the research group, which provided a clear understanding of how the continuous variables interact, sustained by kinetic investigations that had been previously performed.[107] Furthermore, the regioselective outcome of the reaction can be significantly influenced by the choice of solvent, a topic extensively covered in existing literature. This body of literature serves as supporting evidence for the results observed during the optimisation process, further substantiating the achieved reaction outcomes.[108–110]

2.4 Simulated Optimisations

2.4.1 S_NAr Simulation

As previously mentioned, conducting simulated tasks to complement experimental studies holds significant importance. During the COVID-19 pandemic, these methodologies were further highlighted as a necessity for providing optimisation data during laboratory closures. It was during this phase that efforts were initiated to formulate simulated investigations for the upcoming S_NAr project. Initially, the focus was centred on crafting a simulation script aimed at optimising continuous variables in the context of this case study. This choice was guided by the availability of kinetic data specific to the desired S_NAr reaction that had been published by Hone *et al*.[107]

The simulated studies were executed utilising two distinct categories of global optimisation algorithms, which are as follows:

1) SNOBFIT – Designed to facilitate simulated studies concentrating exclusively on continuous variables. This entailed the creation and training of the simulation scripts on MATLAB.
2) Genetic Algorithm – The MATLAB Inbuilt algorithm was employed to optimise simulated studies involving both continuous and discrete variables. This covered studies on both single-objective and multi-objective.

Through numerical analysis of ODEs, the simulation of concentration profiles expressed as $^{dC_A}/_{dt}$, was successfully resolved. Here, C_A symbolises the concentration of a species existing within the system. This process was repeated iteratively for all species present within the system. By utilising Arrhenius kinetics, it becomes feasible to compute the rate constants corresponding to each transformation within the given reaction pathway. This computation draws on data that has been documented previously within the literature. The rate expression can be given by:

$$k = Ae^{-{}^{E_a}/_{RT}} \quad \textbf{(9)}$$

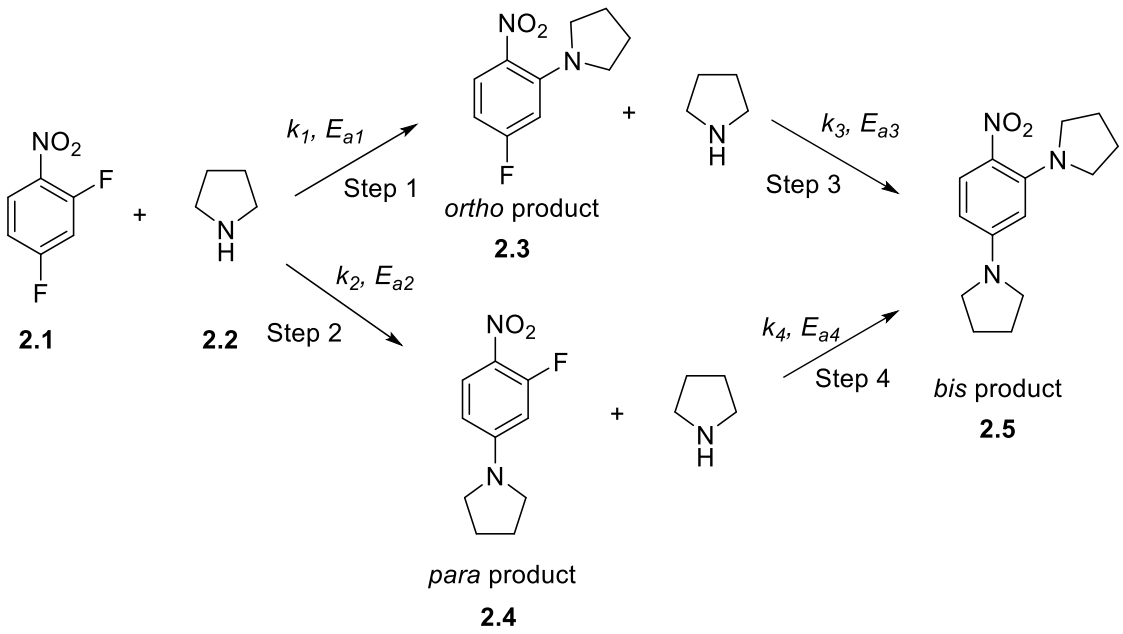
Where, A = Pre-exponential factor, (Units SI)

E_a = Activation energy, (J/mol)

R = $8.3145\ J \cdot mol^{-1} \cdot K^{-1}$ (Ideal gas constant)

T = Temperature, (K), the temperature was later indexed as a variable in the optimisation function.

Using [Eq (9)], the reported pre-exponential factors and activation energies for each step in the reaction pathways, enabling the concentration change model to be built and utilised within an optimisation script. The reaction pathway shown below (Scheme 7) for the S_NAr reaction was the pathway selected for the simulated work and as stated previously would be the reaction pathway selected for the experimental studies to be performed on.



Scheme 7. Kinetic reaction pathway for the formation of all regioisomers in the S_NAr pathway.

For the ODE solver to work effectively for each case study, MATLAB offers different settings to allow for fine-tuning of the solver, the most important settings that can be set for modelling kinetic reactions are:

- NonNegative – Ensures that the concentration of each of the species within the reaction pathway cannot become negative, which is true in real-world chemical systems.
- RelTol – Adjusts the Relative tolerance of the solver which helps in increasing the accuracy of the results but in turn increases the time required. However, moderating the step sizes taken can aid in reducing this time requirement.

With these option settings in place, the subsequent code can be employed to configure the ODE solver, yielding TimeData and ConcData for utilisation in the objective equations. This approach establishes the ODE to solve the reactor model outlined by [Eq ((**10**)-(**14**))]. The resultant simulated concentration profiles can then be harnessed to compute an objective metric; in this instance, the yield of a particular species can be determined based on the concentration data.

A comparative analysis was carried out between two prevalent ODE solvers: ODE45 and ODE15s, the former being a non-stiff solver while the

latter being a stiff solver. In this context, ODE45 showcased superior performance in both accuracy and timing (evaluated using MATLAB's tic toc measurement) for the conducted simulated experiments. This observation can be attributed to the fact that the equations being solved lack terms that could induce rapid fluctuations in the solution, thereby rendering the ODEs in this case non-stiff. A screenshot of the subsequent ODE solver code is presented in Chapter 6.3.3 Figure 52.

$$\frac{dC_A}{dt} = -k_1 \cdot C_A \cdot C_B - k_2 \cdot C_A \cdot C_B \quad \textbf{(10)}$$

$$\frac{dC_B}{dt} = -k_1 \cdot C_A \cdot C_B - k_2 \cdot C_A \cdot C_B - k_3 \cdot C_B \cdot C_C - k_4 \cdot C_B \cdot C_D \quad \textbf{(11)}$$

$$\frac{dC_C}{dt} = k_1 \cdot C_A \cdot C_B - k_3 \cdot C_B \cdot C_C \quad \textbf{(12)}$$

$$\frac{dC_D}{dt} = k_2 \cdot C_A \cdot C_B - k_4 \cdot C_B \cdot C_D \quad \textbf{(13)}$$

$$\frac{dC_E}{dt} = k_3 \cdot C_B \cdot C_C + k_4 \cdot C_B \cdot C_D \quad \textbf{(14)}$$

Where C_A= Concentration of 2,4-difluoronitrobenzene, (M)

C_B = Concentration of morpholine (or pyrrolidine), (M)

C_C = Concentration of *ortho*, (M)

C_D = Concentration of *para*, (M)

C_E = Concentration of *bis*, (M)

t = time, (min)

k_1 = Rate of reaction for the first step, (min^{-1})

k_2 = Rate of reaction for the second step, (min^{-1})

k_3 = Rate of reaction for the third step, (min^{-1})

k_4 = Rate of reaction for the fourth step, (min^{-1})

Using this, it was possible to build a suitable kinetic model to perform simulated experiments on and subsequently optimise using one of the previously mentioned algorithms for their different purposes. This system

proved to be an ideal example for simulation since the discrete variable optimisation would be performed on the same case study where differentiation of the solvent would suitably affect the regioselectivity outcome of the reaction.

For the simulation of the S_NAr reaction, only continuous variables were studied i.e., temperature, residence time, concentration of 2,4-difluoronitrobenzene (**2.1**) and the equivalents of the pyrrolidine/morpholine (**2.2**) due to the publication of the kinetic data for only one of the solvents, with that being ethanol. Boundaries were set for the continuous variables such that, the upper and lower limits would fit the experimental system's capabilities and would be selected to be the limits for the experimental study. Different objectives such as yield, productivity and E-factor were all explored for different parts of the simulated study, the yield of different products being the main objective for each study due to the nature of the regioselective outcome for the discrete variables in the future experimental study. The boundaries for the simulated optimisation are shown in Table 3.

Table 3. Boundaries for continuous variables in the S_NAr simulation with morpholine.

	Residence time (min)	*Temperature (°C)*	*SM (**2.1**) concentration in reactor*	*Equivalents of morpholine (**2.2**)*
Minimum bounds	0.5	60	0.1	1
Maximum bounds	2	120	0.5	5

In the initial simulation, SNOBFIT was selected as the algorithm of choice and the *ortho* (**2.3**) – product yield was optimised in an experimental run followed by the *para* (**2.4**) – product due to the nature of SNOBFIT being a single objective optimisation algorithm. Both sets of simulations were performed initially with the pyrrolidine kinetic dataset followed up by the morpholine dataset, which would allow for a direct comparison between the simulated and experimental optimisations.

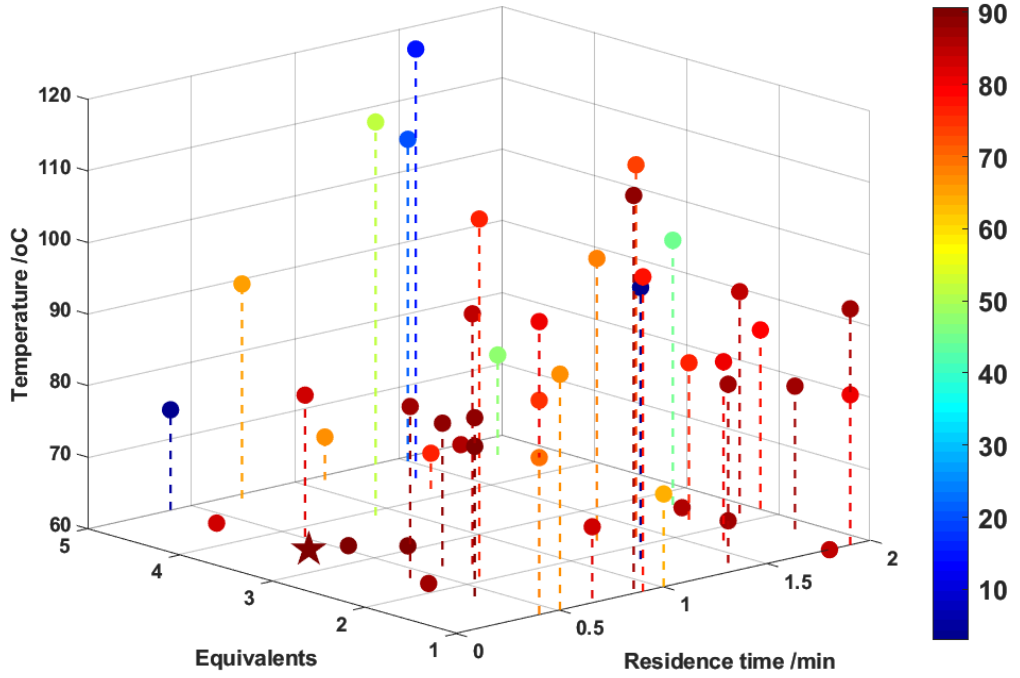


Figure 18. SNOBFIT simulation for the S_NAr reaction with pyrrolidine with the *ortho* product selected as the objective. The colour bar for this graph represents the yield for the *ortho* product, with a range of 0-100%. Where the ★ represents the optimum *ortho* yield.

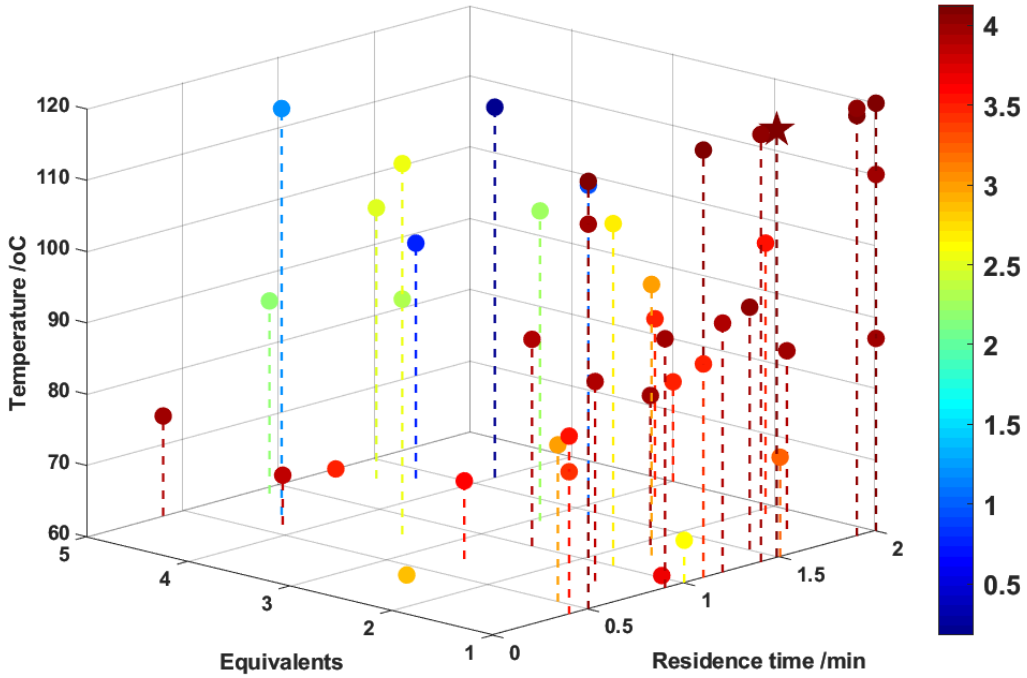


Figure 19. SNOBFIT simulation for the S_NAr reaction with pyrrolidine with the *para* product selected as the objective. The colour bar for this graph represents the yield for the *para* product, with a range of 0-4%. Where the ★ represents the optimum *para* yield.

The results show that in the simulation the algorithm effectively explores the design space to find the global optimum value of 90% *ortho* (**2.3**) yield, which is represented by the square symbol in Figure 18. This initial dataset highlights the effectiveness of the SNOBFIT algorithm in identifying the global optimum in the optimisation space for these simulated studies. The conditions required to form the maximum *ortho* (**2.3**) yield, utilised low temperatures and equivalents to prevent overreaction from the *ortho* (**2.3**) to the *bis* (**2.5**) product. From this information, another simulation was performed using the model for the morpholine reactant to compare the outcome of the S_NAr reaction between each nucleophilic species. The simulation of the morpholine species was especially important as it would be the nucleophile of choice when performing the solvent optimisation following the conditions reported by Schweidtmann *et al.*[66]

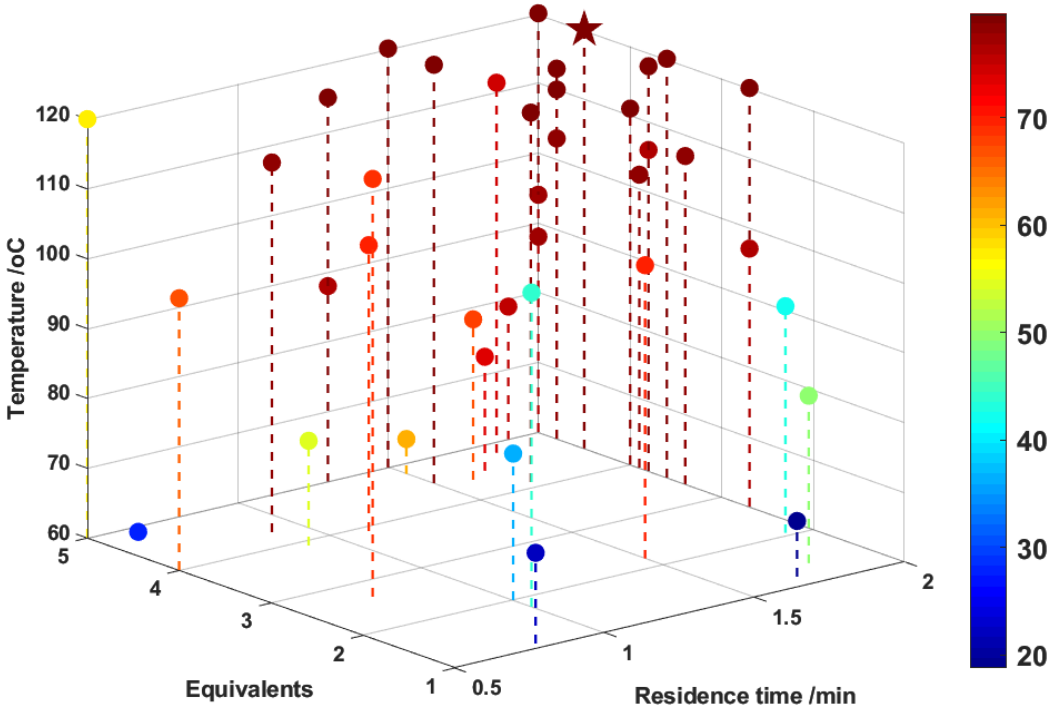


Figure 20. SNOBFIT simulation for the S_NAr reaction with morpholine where the *ortho* (**2.3**) yield is selected as the objective. Where the ★ represents the optimum *ortho* yield.

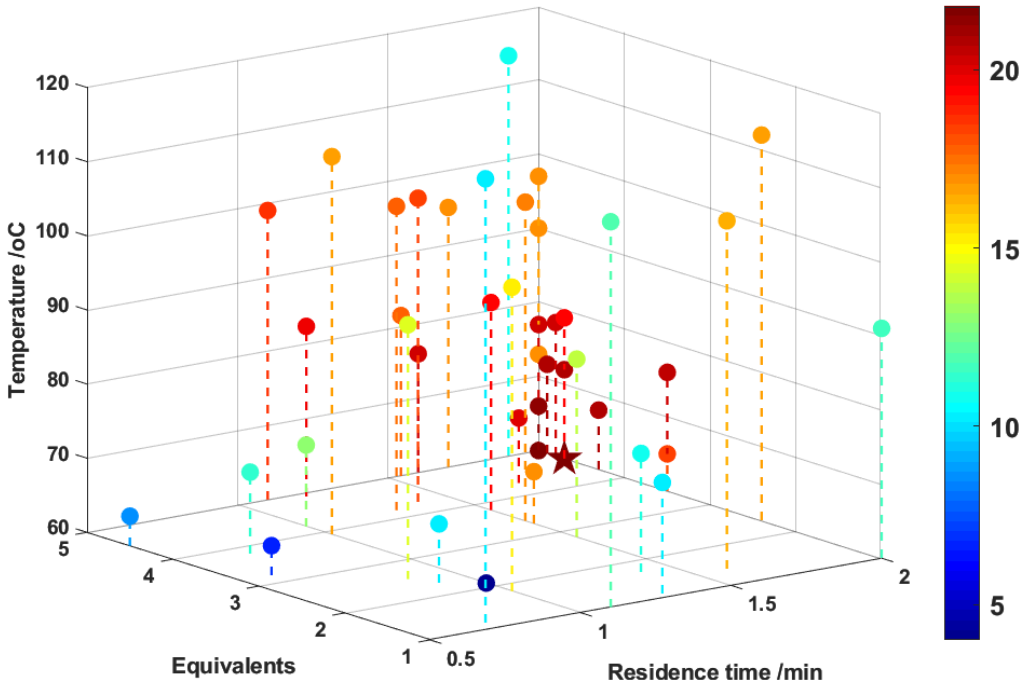


Figure 21. SNOBFIT simulation for the S_NAr reaction with morpholine where the *para* (**2.4**) yield is selected as the objective. Where the ★ represents the optimum *para* yield.

Interestingly, when morpholine is used in replacement of the pyrrolidine, it causes the maximum yield of the *ortho* product to decrease, where the maximum amount of *ortho* formed is 79.7% as highlighted in the red square of Figure 20. When performing the simulated optimisation on the *para* product, this decrease in *ortho* yield can be observed with an increase in the *para* yield from 4% with pyrrolidine up to 22% with morpholine, which led to further rationality for the selection of morpholine for the experimental studies. This change in reactivity can be linked to steric effects between the nitro group and the morpholine (bulkier than pyrrolidine) such that less *ortho* is formed. These results go on to justify the experimental observations of acetonitrile outperforming methanol when maximising the *ortho* product formation.

Following the initial investigations, work commenced on performing a multi-objective simulated optimisation on the same case study to delve into the optimisation of multiple objectives for the S_NAr reaction, an area that had not been extensively reviewed. Utilising the MATLAB genetic algorithm, 31,000 iterations were performed to evaluate and optimise the two selected objectives. While productivity and overall yield were initially chosen as the objectives to maximise, the focus later shifted to studying the yields of the

ortho and *para* products to align with the ongoing experimental work (refer to Section 4.2.1). The simulated work utilised morpholine kinetic data, consistent with the species employed in the experimental solvent optimisation process, ensuring coherence between the simulation and experimental studies.

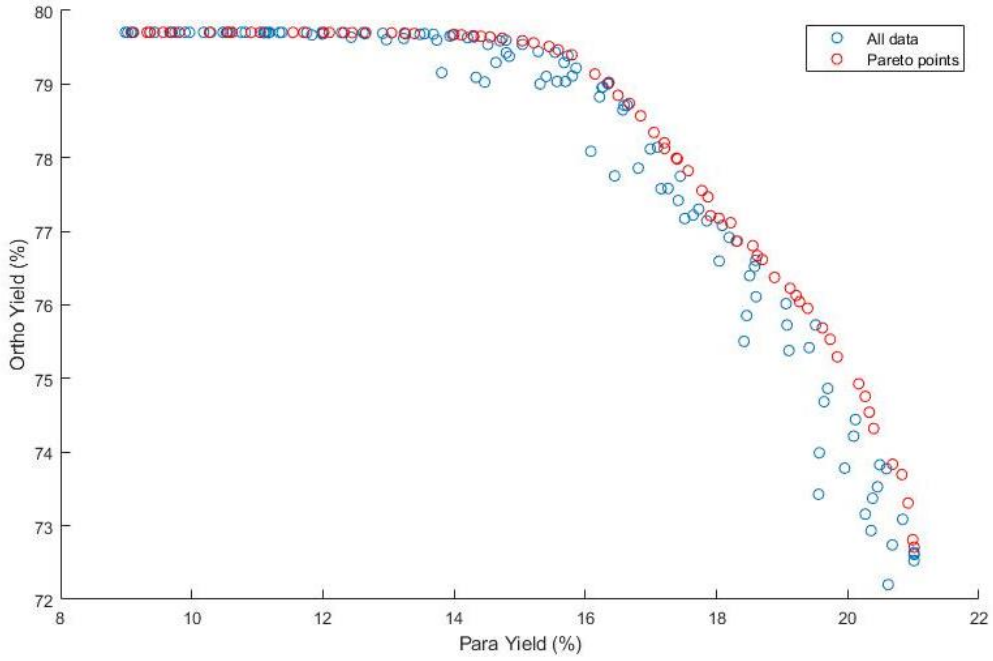


Figure 22. Illustrates the results of a multi-objective optimisation approach using a genetic algorithm to optimise the S_NAr reaction. The objectives of this optimisation were to maximise the yields of the *ortho* (**2.3**) and *para* (**2.4**) products from the reaction.

As anticipated, the simulation effectively highlighted the trade-offs between the formation of each of the regioisomers under study. The optimisation revealed that low temperatures were essential to maximise the yield of either the *ortho* or *para* products, emphasising the sensitivity of the S_NAr reaction for overreaction to form the *bis* product due to temperature variations.

From the 31,000 data points generated, a total of 200 Pareto-optimal solutions populated the Pareto front. These solutions represent the optimal trade-offs between the two selected objectives and are visually distinguished by highlighting the points in red in Figure 22. The Pareto front provides valuable insights into the relationships and dependencies between the yields

of the *ortho* and *para* produces, aided in guiding further experimental design and optimisation strategies for the experimental solvent study.

Performing these simulated optimisations using the experimental kinetic data gave a key insight into the effects that the continuous variables could have on the output of the reaction and explained some of the reactivity later observed in the experimental work for the ethanol solvent. This simulated work allowed for the refinement of the experimental process and provided key justification for the results observed with the ethanol solvent. Finally, it allowed for 'experimental' work to be continued to develop the project throughout the COVID-19 pandemic and provided key training and further enhancement of coding skills.

2.5 Experimentally Optimised Results

Using the simulated optimisation provided key insight into the proposed S_NAr solvent optimisation, although performed on only a singular solvent, it acted as a benchmark for the choice of the solvents being selected for the optimisation. As it is known in the literature, the solvent properties can significantly affect the regioselective outcome of the reaction.[108,109,111] For this optimisation three common solvent polarity metrics were evaluated: (i) polarity index, which is a measure of a solvent's ability to interact with various polar test solutes; (ii) dipole moment, which is calculated based on the product of the magnitude of separated charges and the distance between those charges; (iii) dielectric constant, which is a measure of a substance's ability to insulate charges from one another, Table 4. Solvents were also selected based on these polarity attributes in addition to being able to provide a homogenous reaction mixture.

Table 4. Chemical descriptor values of common solvent polarity metrics.[112]

Solvent	*Polarity Index*[a]	*Dipole Moment*	*Dielectric Constant*
NMP	6.7	4.09	32.20
DMAc	6.5	3.72	37.78
DMF	6.4	3.86	36.71
MeCN	5.8	3.44	37.50
EtOH	5.2	1.66	24.55

Due to the regioselective outcome of the reaction, and to further highlight the solvent effects on the objective outcomes, the yield of both the *ortho* (**2.3**) and *para* (**2.4**) regio-isomers were selected as the process metrics to be examined. Given that MVMOO is a minimisation algorithm, the negative response of both objectives was input to achieve objective maximisation, leading to the formulation of the optimisation equation [Eq (15)]. The continuous variables for the optimisation procedure were performed on four variables residence time (mins), starting material (**2.1**) concentration (M), pyrrolidine (**2.2**) equivalents and temperature (°C).

$$minimise\ [-(yield_{ortho}), -(yield_{para})] \quad \textbf{(15)}$$

Subject to:

$$Solvent \in [NMP, DMAc, DMF, MeCN, EtOH]$$

$$residence\ time\ \in [0.5, 2.0]$$

$$Concentration\ (\mathbf{2.1})\ \in [0.05, 0.175]$$

$$Equivalents\ (\mathbf{2.2}) \in [1, 5]$$

$$Temperature\ \in [60, 120]$$

To achieve switching of the discrete variable stock solution, a 7-port 6-way switching valve was implemented with the feeds from each discrete variable stock solution taking positions 1-5. The final 6th slot was left free for a solvent flush between experiments, this would ensure that the outlet line from

the switching valve would be cleaned between each iteration and minimise cross-contamination between discrete variables. For this S_NAr case study, a precautionary mixture of DMSO and water (5% *v/v* water in DMSO) was used, as in preliminary exploration experiments, it had been noted that solid formation was occurring at higher temperatures when NMP was selected for the variable of choice. After further research, it highlighted that NMP had poor solubility of triethylamine salts, causing a build-up in the reactor.[113] This led to the introduction of the DMSO/water washing solvent, where water was added to the solvent to ensure that the solution did not freeze overnight in the lab. In addition, the potential formation of these salts led to a reduction in the 2,4-difluoronitrobenzene concentration that had been previously reported in the paper, to the concentration presented in [Eq(15)]. The finalised schematic which was used throughout the experimental optimisation of the S_NAr reaction is provided in Figure 23.

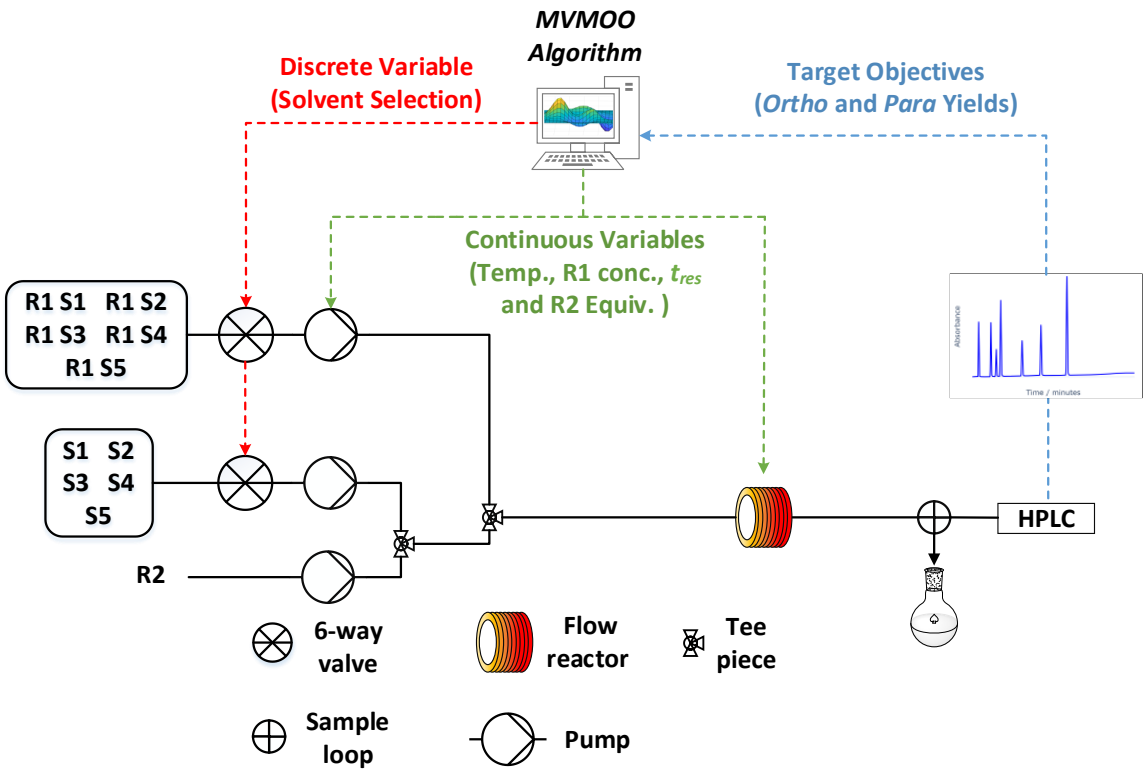


Figure 23. Automated flow platform utilised within the solvent optimisation of the S_NAr reaction, where *ortho* and *para* yields are selected as the objective for the optimisation.

Initialisation of the MVMOO algorithm began with a 5 sample LHC where each solvent was afforded 5 experiments resulting in an initial 25 experiments. The algorithm was allowed to run for a further 74 experimental iterations to effectively map out the Pareto front for the system. Of these experiments, a total of 20 of them were identified as non-dominated solutions which sufficiently highlighted the trade-off between the formation of the *ortho*-(**2.3**) and *para*-(**2.4**) regioisomers. In this case, the Pareto front can be observed to be broken into three sections, resulting from the different solvent effects being observed in the system, Figure 24. The optimal *ortho*-(**2.3**) yield was achieved by MeCN and corresponded to an 80% yield which related to a 10% *para*-(**2.4**) yield. Conversely, the optimal *para*-(**2.4**) yield was achieved by NMP and corresponded to a 48% yield which related to a 49% *ortho*-(**2.3**) yield.

DMF was unique in providing a moderate compromise to the formation of both objectives under study to result in dominating the mid-section of the Pareto front. Despite the structural similarities between DMF and DMAc, it could be seen that DMAc gave a much higher selectivity towards *para*-(**2.4**) than its structural counterpart, this can be related to DMAc obtaining a higher polarity index which in turn aids the selectivity. At the top, for the highest yielding *para* results, NMP dominated in all experiments apart from one. On the other side of the Pareto front, although EtOH provides higher *ortho*-(**2.3**) selectivity, it suffers from forming significantly lower overall yields resulting in MeCN dominating this Pareto section. Additionally, it resulted in the MVMOO algorithm only suggesting a singular EtOH-based point after the initial LHC which can be seen to perform extremely poorly leading to the algorithm focussing on MeCN for this region.

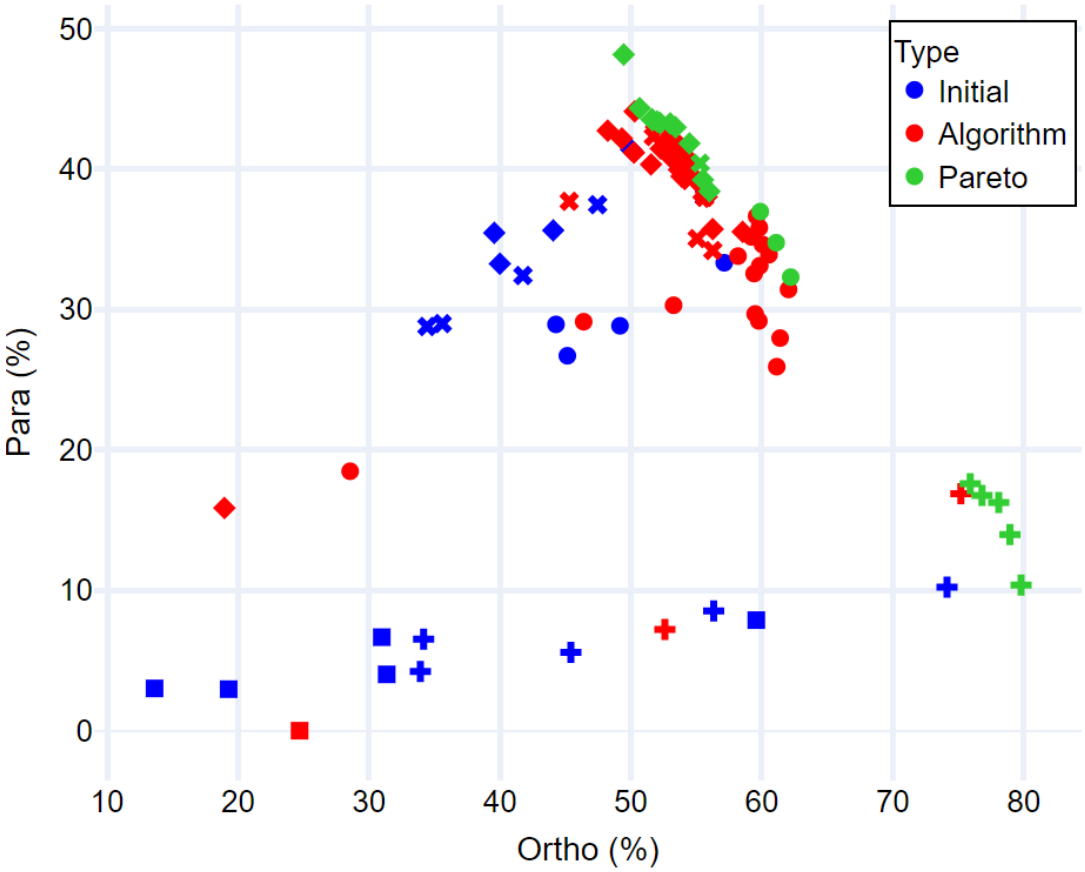


Figure 24. Results of the five-parameter mixed variable optimisation on the S_NAr reaction, where an initial 25 experiment LHC was conducted with a further 74 additional iterations identifying 20 non-dominated solutions. Solvent shapes represent: ● – DMF, ◆ – NMP, ■ – EtOH, ✚ – MeCN, ✖ – DMAc.

It can be observed how the MVMOO algorithm was able to effectively map out the relative importance of the continuous variables combined with discrete variables across the Pareto front, Figure 25 and Figure 26. To maximise the *ortho*-(**2.3**) yield, when using MeCN it requires high residence times, high equivalents and temperatures whereas the concentration of 2,4-difluoronitrobenzene (**2.1**) is a less important variable in this case. On the other hand, for NMP to achieve higher yielding *para*-(**2.4**) results, it could be achieved over a range of residence times, equivalents, concentrations, and temperatures indicating how strong the solvent effects are for these results.

Similarly, to MeCN, when observing the results from DMF, to achieve a moderate compromise between objectives, it required higher equivalents, temperatures, and concentration. The contrasting results for each discrete variable highlight the importance of these types of workflows, as these

observations are contradictory to the assumptions of traditional optimisation procedures. In these types of methodologies, the continuous variables are optimised independently of the discrete variables, where it would be wrongly assumed that the interactions between the discrete and continuous parameters are the same throughout for each discrete variable under study. The MVMOO algorithm was able to successfully optimise the mixed variables simultaneously, providing a greater understanding of the interactions of the variables to explore the trade-off of the objectives. Additionally, traditional methods for optimising continuous and discrete variables sequentially can lead to elongated timeframes for the full study of the chemical systems, requiring material and time consumption to be increased.

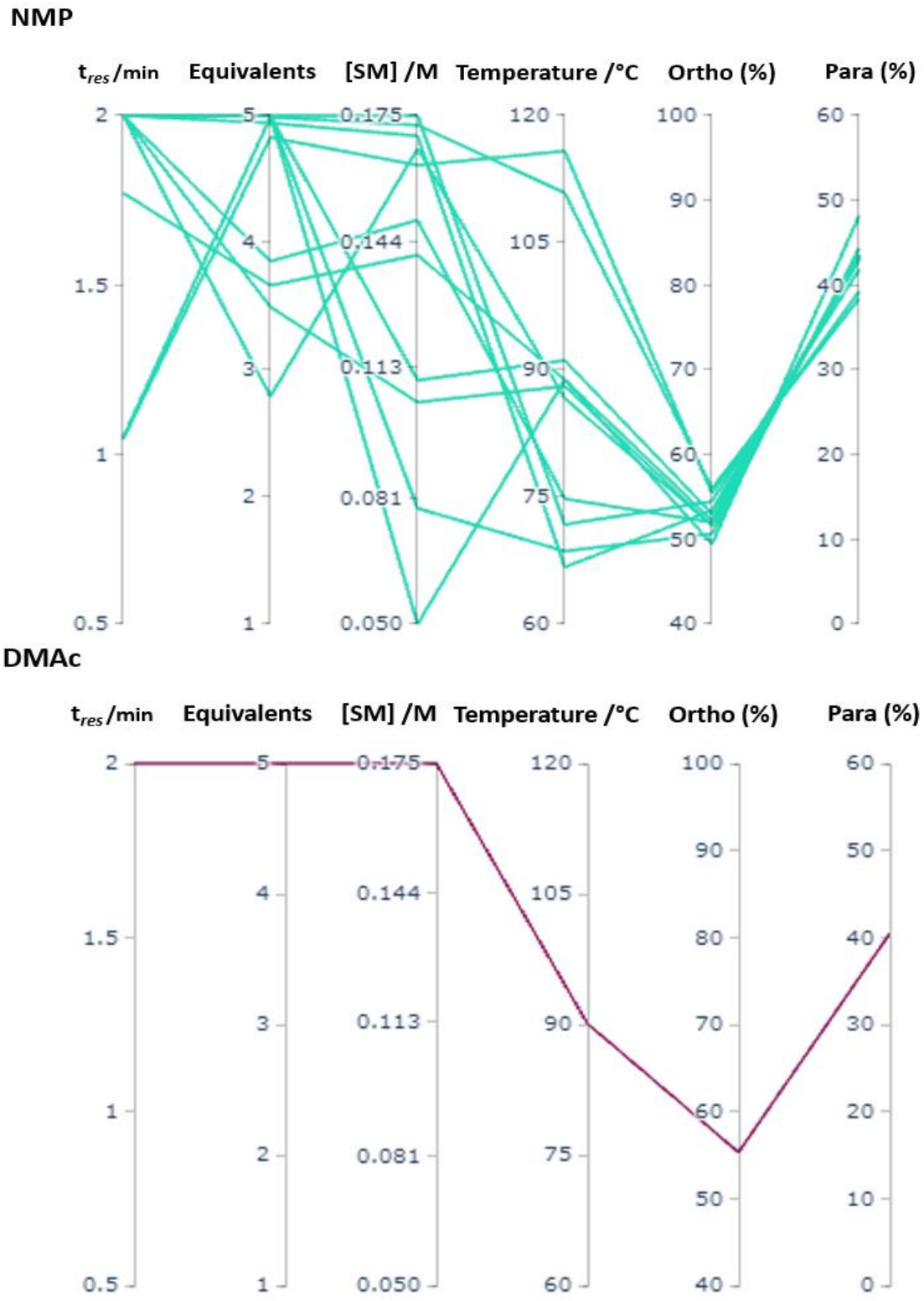
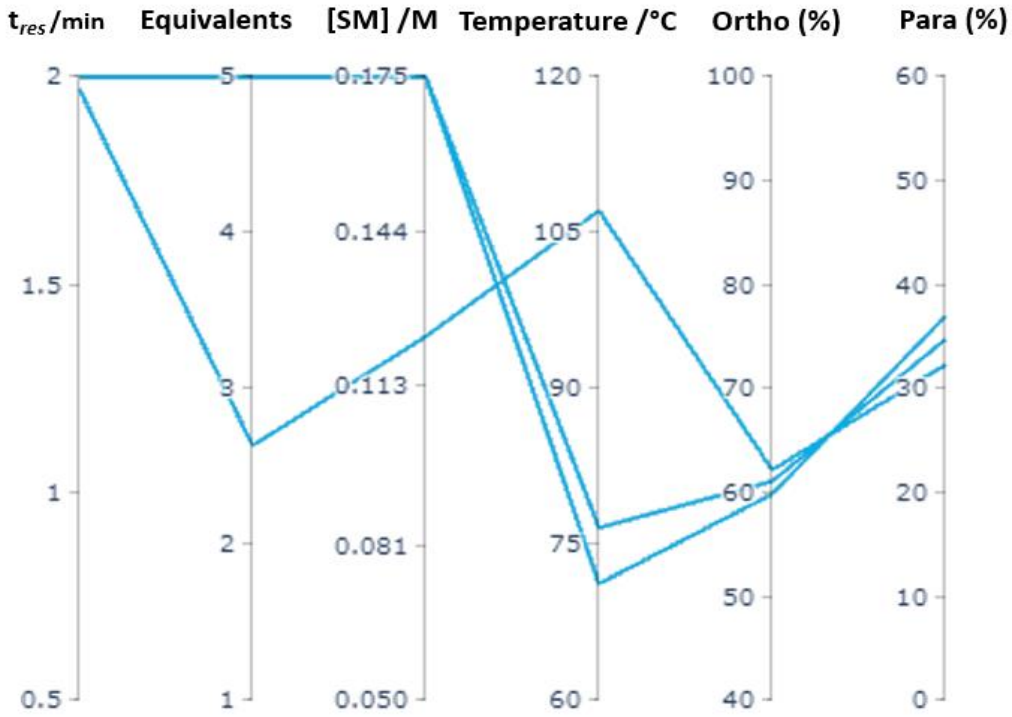


Figure 25. Parallel coordinate plots for NMP and DMAc for the Pareto optimal solutions of each discrete variable highlighting the required interactions between the four continuous variables for each non-dominated solution. The error for % yield of both *ortho* and *para* was calculated using 10 repeat experiments from the optimisation, where an error of $\pm$4% was observed for *ortho* % yields and an error of $\pm$3% was observed for *para* yields.

DMF



MeCN

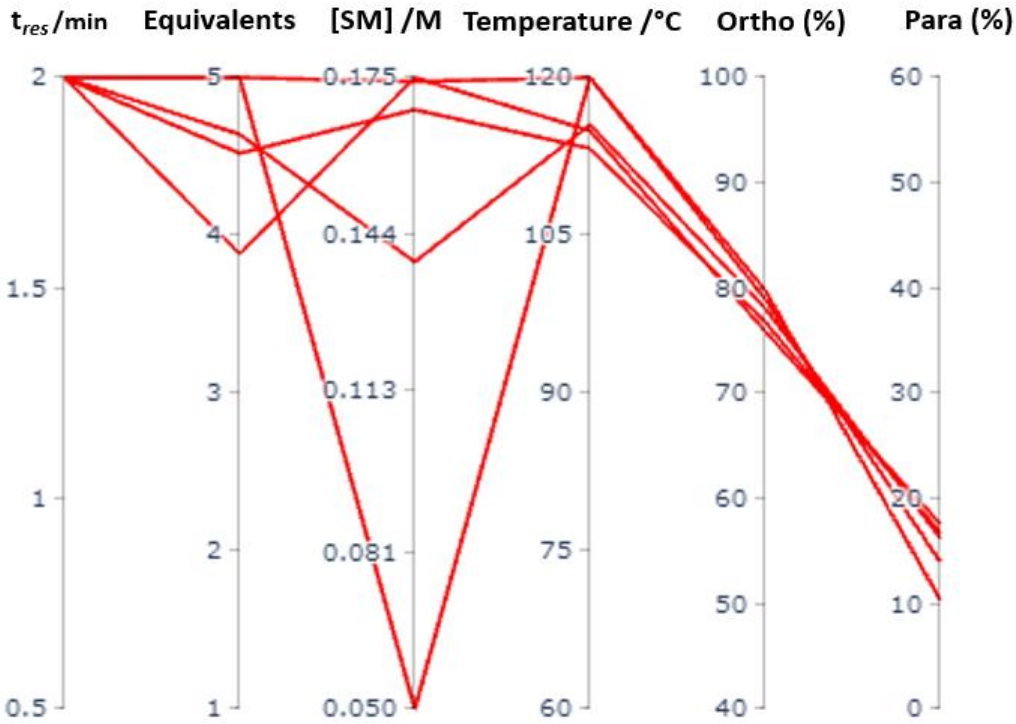


Figure 26. Parallel coordinate plots for DMF an MeCN for each Pareto optimal solution produced by these solvents. The error for % yield of both *ortho* and *para* was calculated using 10 repeat experiments from the optimisation, where an error of ±4% was observed for *ortho* % yields and an error of ±3% was observed for *para* yields.

Employing MVMOO enabled this optimisation to be completed in only 18 hours, requiring no human intervention, thus highlighting the increased efficiency of this process when compared to iterative HTS methodologies or traditional sequential optimisations for each solvent.[99] Furthermore, commissioning MVMOO for the simultaneous optimisation of mixed variables, allows for real-time monitoring of the hypervolume as the experiments proceed. This acted as a key part of information for the stopping criteria of the experiment, as it was decided that once the optimisation had reached a minimum of 60 experiments, the hypervolume would be monitored until it reached a plateau that lasted for a duration of 5 experiments.

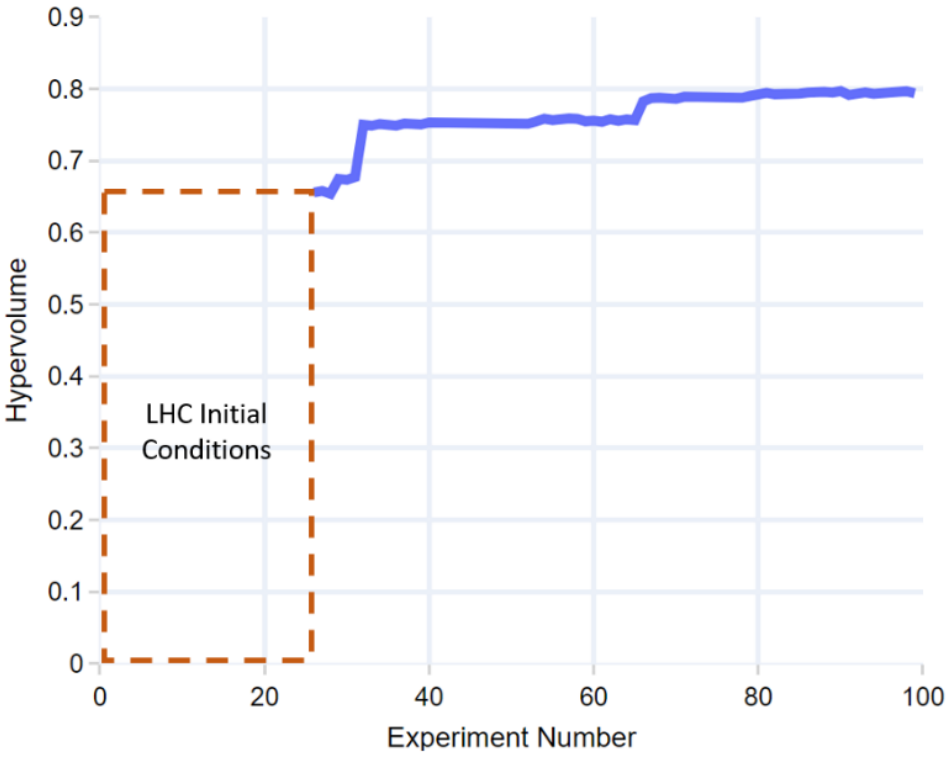


Figure 27. Hypervolume plot for the S_NAr reaction that was monitored throughout the optimisation procedure. After 60 experiments monitoring enabled the termination at 100 experiments once a plateau was achieved.

As can be seen in Figure 27, after 60 experiments, there is a gradual gain in hypervolume as the optimisation proceeded until around 90 experiments where it can be noted that a plateau was reached, leading to the finishing of the optimisation. Doing it in this manner, allowed for the conservation of material and time, in addition to gaining a further 9 Pareto front points.

Using the dataset gained from this optimisation, Jamie Manson went on to complete some key chemical descriptor analysis to observe the underlying solvent polarity metric effects on the regioselective outcome. Herein, the key information related to these studies will be reported, however, it should be noted that all simulated data and graphical representations were completed by Jamie Manson. Overall, from the chemical descriptor analysis, it can be observed that there is a general bias towards the formation of the *ortho*-(**2.3**) regioisomer, which is likely related to the hydrogen bonding between the nucleophile and the nitro group. The models deduced that increasing the *para*-(**2.4**) regioisomer yields was sufficiently correlated to the increase of the polarity index of the solvent, Figure 28. Furthermore, the studies contributed to further insights into the continuous variables' relationship with forming each regioisomer. These results provided further justification for the hypothesised relationships that had been identified by the MVMOO algorithm on the chemical system throughout the optimisation.

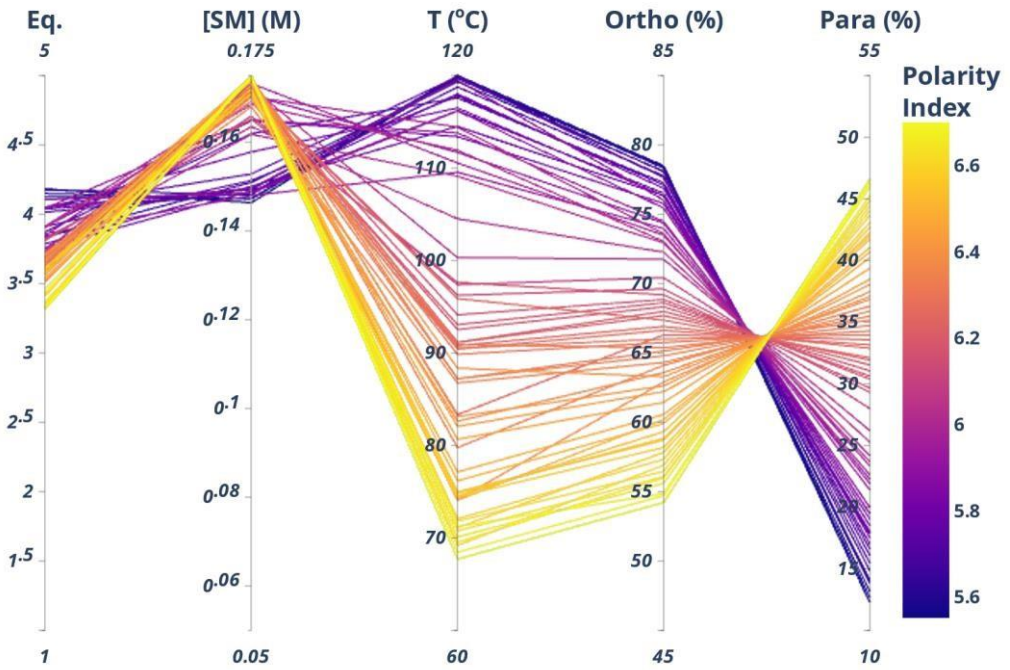


Figure 28. Parallel coordinates plot for the simulated Pareto front using polarity index as an input continuous variable. Residence time was maintained at 2 minutes and thus excluded from the figure. Simulation and figures were produced by Jamie Manson.

Conversely, during the chemical descriptor analysis, when selecting either the dipole moment or dielectric constant as inputs, it failed to effectively

illustrate the observed relationship. When inputting the dielectric constant, it predicted DMAc and MeCN to perform in a similar manner, which is contrary to the experimental results that are observed to be the complete opposite. The use of these studies can significantly highlight an important area of algorithm application with additional data extraction for potential use in process optimisation and exploring chemical spaces. It enables the ability to suggest ranges of conditions that had not been performed within the experimental procedure and explore further relationships. However, it should be noted that this type of methodology was only completed on a small number of solvents to identify the optimal ones, where a much greater dataset would be required to examine the key solvent descriptors in further detail, due to the vast other solvent properties that could be in effect. Additionally, to be able to utilise solvents in the suggested polarity range from the simulation, it would require solubility studies to be performed to optimisations, in which there is no consideration performed by the model.

2.6 Conclusion

The combination of automated flow processing with intelligent algorithms represents a key area of research required to improve process efficiency and reduce the material consumption of optimisations. Consideration of both discrete and continuous variables in a simultaneous manner remains a relatively unexplored area, where traditional workflows employ sequential optimisations to incompletely explore the interactions. Early-stage reaction development and screening is a key area where the application for simultaneous optimisation can have a significant effect.

In this chapter, the work is based on employing a mixed variable optimisation algorithm onto an automated flow platform, studying the solvent effects in combination with continuous variables on the outcomes of the S_NAr reaction. To further study the regioselective effects of the solvents, the work presented utilised both the desired *ortho*- and *para*-yields as the objectives to be studied in a multi-objective format. The workflow used a four continuous variable system in combination with 5 solvents as the discrete variables to elucidate the variable-variable interactions required to explore the trade-off

between the objectives. Before this optimisation, using published kinetic data for the system, simulations for one of the solvents (EtOH) were completed to increase prior knowledge of the system and allow for fine-tuning of the continuous variables to better suit the experimental optimisation. Furthermore, these initial studies provided support for the observation that MeCN outperformed EtOH in producing non-dominated solutions with respect to *ortho* yields within the allotted experiments. Finally, chemical descriptor analysis by Jamie Manson using the optimisation data highlighted the significant importance of the polarity index as a chemical descriptor on the regioselective outcome of the reaction.

This work represents the first completed studies of exploring multi-objectives for mixed variables on a continuous flow chemical system, offering enhanced efficiency and identification of key interactions over traditional sequential optimisation workflows. In addition, performing these studies on an automated flow system offers the additional benefits of access to higher temperatures and hazardous intermediates when compared to mixed variable HTS batch systems. The enhanced efficiency of these types of optimisations lends itself towards early-stage reaction development, where there is a desire to maximise information gained and a reduction in material and time consumption.

Chapter 3. Exploring API Synthesis: A Mixed Variable Optimisation Approach of Catalytic Systems

3.1 Introduction

Palladium-catalysed cross-coupling reactions have transformed the field of organic chemistry and the pharmaceutical industry alike, providing enhanced access to the synthesis of natural complex structures.[114,115] Over the years, these types of catalysed reactions have received extensive research with great success in developing a vast scope of possible reactions utilising a substantial number of different variants of catalysts, ligands, and additives to make the reactions possible. However, these types of reactions suffer from the catalyst itself, due to palladium being a precious metal, it can lead to staggering prices for the catalysts, requiring research of other metal catalysts such as nickel and copper to act as replacements.[116] Furthermore, the desire to reduce catalytic consumption has led to substantial studies and significant interest from the pharmaceutical industry in the adoption of better optimising these chemical systems as well as the recycling of the palladium catalysts.

Within the ever-growing pharmaceutical landscape, the impact and effect the chemical industry is having on the environment is an increasing concern, leading to companies such as GSK in Singapore and Eli Lilly adopting greener techniques for chemical synthesis.[117,118] The pharmaceutical industry is no exception, with many attempting to minimise the effect it is having, whether adopting new waste minimisation strategies or recycling solvents to reduce their impact. The implementation of new technologies such as continuous flow processing has enabled the adoption of greener methods for API synthesis within the pharmaceutical industry. Integration of continuous flow manufacturing, combined with other alternative chemical methodologies e.g., photo flow chemistry, and electro flow chemistry can facilitate access to new chemistry in flow processes at novel research level but also large-scale production.[119,120] All of which provide the added benefits of flow processes over traditional batch methodologies, further enhancing the green aspect of these systems. Furthermore, the adoption of machine learning algorithms to optimise these chemical systems can further aid in the green aspects of the processes, helping enhance the energy efficiency and preventing any wasted reactions or materials.[20]

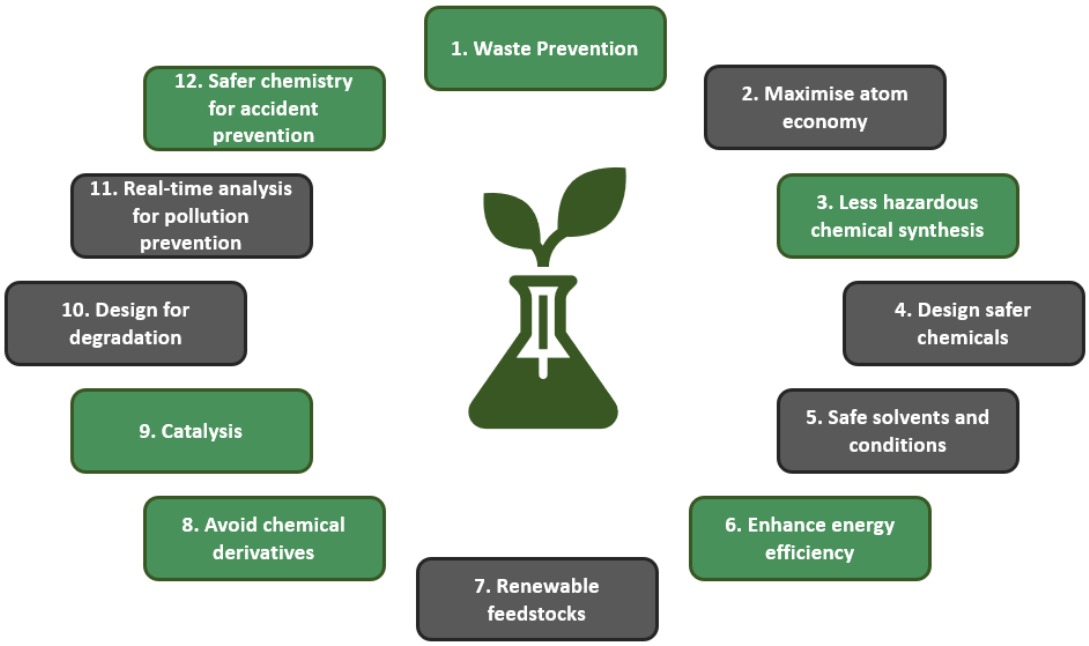


Figure 29. 12 principles of green chemistry with highlighted areas addressed by these types of mixed variable optimisations presented.

Green chemistry and engineering are two highly important fields for addressing and responding to environmental and sustainability challenges stemming from the chemical industry. The 12 principles of green chemistry were developed in 1991 by Paul Anastas and John Warner to provide a useful guideline for other industries to reduce the risk of their processes as well as minimise the environmental footprint.[121,122] The introduction of these principles in addition to the ever-increasing concern of the impact that the chemical industry has on the environment, has led to many modifying traditional techniques to better fit these guidelines with numerous reviews on the benefits of Green Chemistry being published.[122–127] This approach of mixed variable optimisations of catalytic reactions addresses the 12 principles by providing routes to less hazardous chemical synthesis and safer chemistry to reduce accidents coming from performing them under flow conditions. All whilst enhancing energy efficiency and reducing waste coming from less wasted experiments required to complete a successful optimisation when compared to other traditional sequential methods.[128]

The inclusion of catalytic systems in autonomous self-optimisation processes has been explored for both single-objective and multi-objective

optimisations where all the literature solely focuses on the manipulation of continuous variables. The main techniques for the adoption of discrete variables within these catalytic optimisations have come through HTE to screen through vast libraries.[129–131] As previously highlighted in Chapter 2, the HTE approach to optimisation of chemical systems results in incomplete process understanding due to the neglect of the potential interactions between the continuous and discrete variables. Thus, becoming an increasing unattractive approach for these types of optimisations. This holds especially true for case studies involving catalytic transformations, where the complexity of the interactions is significantly increased due to the numerous combinations of catalytic species and ligands available. For example, the sequential optimisation of discrete and continuous variables would mitigate the observation of the effects of the temperature of the reaction on the activity of different catalysts. This heightened complexity highlights the necessity for simultaneous optimisation of mixed variable systems on catalytic systems. Jensen *et al*. explored a series of mixed variable transition metal catalysed cross-coupling reactions employing an optimal DoE based algorithm to conduct and direct the optimisation.[74,79,87]. Notably the authors studied a palladium catalysed Suzuki-Miyaura cross-coupling reaction employing a micro-fluidic system to minimise consumption of materials whilst operating parallel reactions throughout the optimisation.

These approaches utilised an adaptive response surface methodology that iteratively eliminated catalytic species from the optimisation. Whilst these processes presented the efficient optimisation of mixed variable chemical systems on a continuous flow platform, employing a black-box algorithm, they were limited to maximise a single-objective. This type of methodology does not provide understanding into the trade-off (Pareto front) between conflicting performance criteria, which is critical in the development of feasible industrial processes. Furthermore, limiting the system to only single-objectives confines the variables interactions observed, as it only studies the relationships required for optimisation of a single metric.

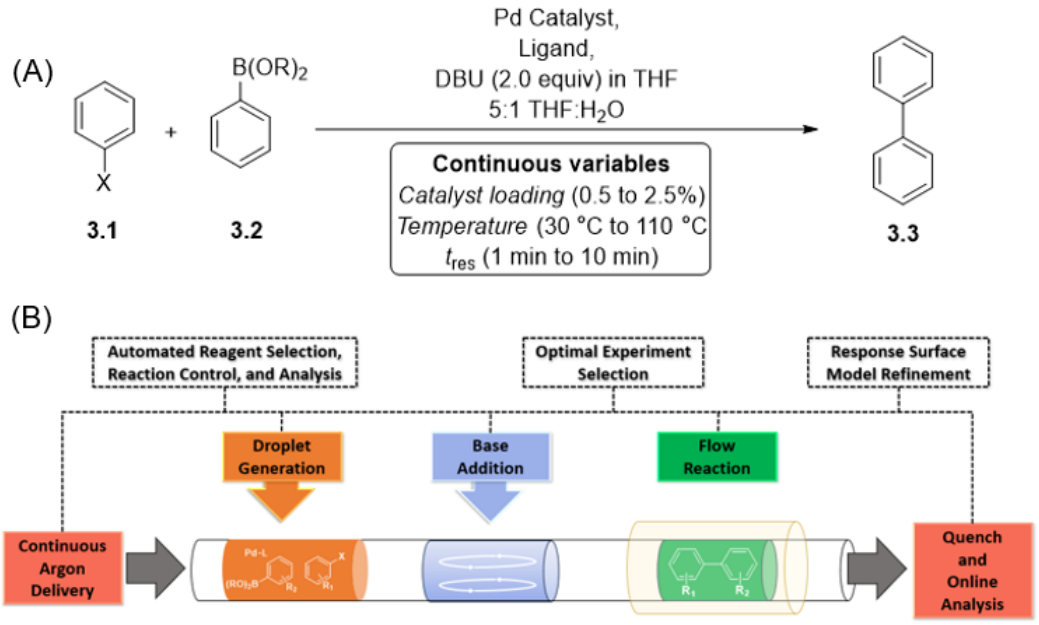


Figure 30. (A) Schematic for the Suzuki-Miyaura coupling reaction optimised with respect to mixed variables on a micro-fluidic platform.

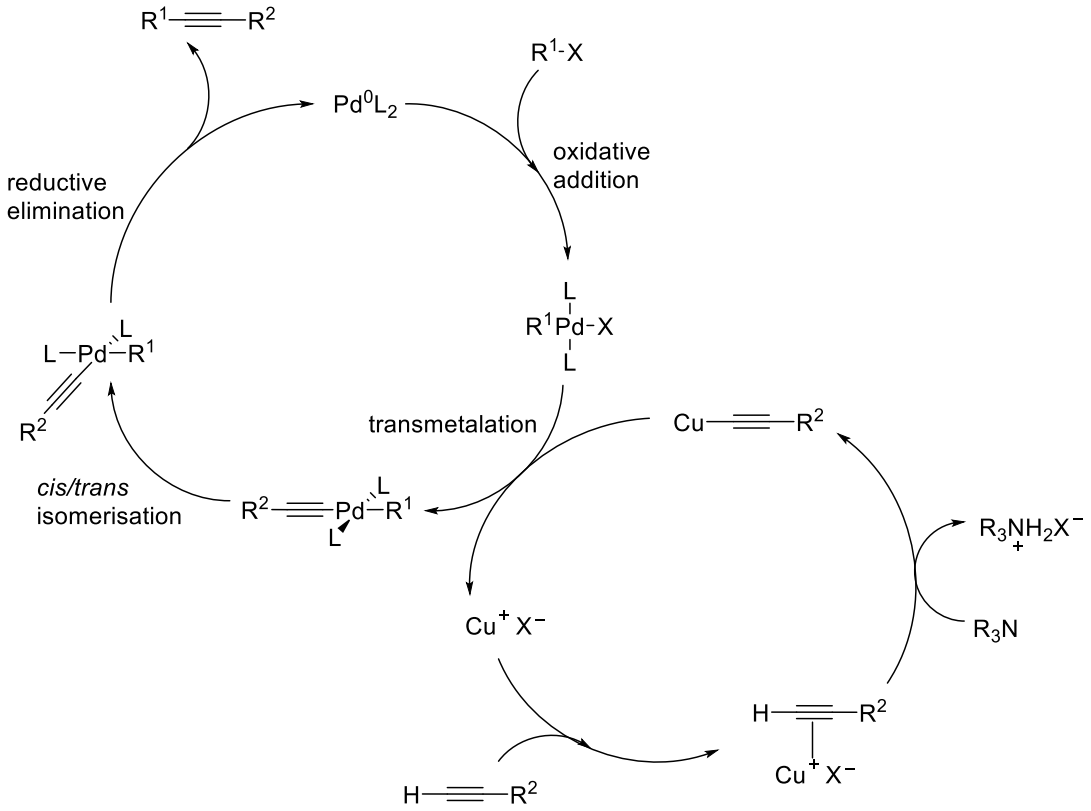
(B) Flow over time transformation of the reaction slug within microfluidic platform that was utilised within the optimisation process.

The case study proposed for this chapter aims to address the identified limitations of previous optimisation in catalytic chemical reactions whilst incorporating mixed-variables into the system for the study of multiple objectives. Inclusion of mixed-variables for the evaluation and optimisation of catalytic systems also aids in increasing the complexity of the chemical reaction present, further extending the value of these types of optimisation processes. This optimisation focussed on maximising the productivity of the reaction outcome whilst minimising the wastage of materials, achieved through maximisation of the reaction mass efficiency. Selection of the discrete variable was made such that the variable of choice would have a significant impact on the catalytic cycle itself and was decided that the ligand choice would be the best fit. This enabled comparisons between high and low-cost ligands which would add an extra consideration of the price to productivity outcomes for such ligands. It is important to highlight that the catalytic system utilised in this study is homogeneous, offering a well-defined environment for investigating the interactions and optimisation of reaction parameters. This systematic approach facilitates the exploration of mutli-objective optimisations and allows for the study of factors related to the catalytic species within the

solution that may influence the objective metrics. Overall, this enhances the understanding of the underlying mechanisms influencing the reaction pathways.

3.2 Sonogashira Cross-Coupling Reaction

The Sonogashira cross-coupling reaction is widely used in organic synthesis for the construction of carbon-carbon bonds between aryl or vinyl halides and terminal alkynes, having been developed in 1975 by Kenkichi Sonogashira. The reaction utilises an active Pd(0) catalyst which can be generated *in-situ* via ligand dissociated or reduced from a precatalytic Pd(II) species. Within the reaction cycle, the process beings with an oxidative addition of the aryl/vinyl halide (R^1-X) to the Pd(0). It is important to highlight that the rate of addition is accelerated when dealing with substrates featuring lower electron density on the C-X bond, following the general trend towards oxidative addition rates: vinyl iodide ≥ vinyl triflate > vinyl bromide > vinyl chloride > aryl iodide > aryl triflate ≥ aryl bromide >> aryl chloride.[132] At the same time, the Cu(I) salt undergoes a side catalytic cycle where the terminal alkyne coordinates, leading to the formation of a π-alkyne-copper complex. This coordination step serves to enhance the reactivity of the acetylenic proton, facilitating its deprotonation in the presence of a base, resulting in the generation of a copper acetylide. Subsequently, through a transmetallation step, the copper acetylide undergoes conversion into the Pd(II) complex to yield the palladium acetylide, regenerating the Cu(I) salt. Finally, following a cis/trans isomerisation process, the reaction is completed with a reductive elimination step. This yields the desired aryl/vinyl alkyne and simultaneously regenerates the active Pd(0) catalyst.[133]

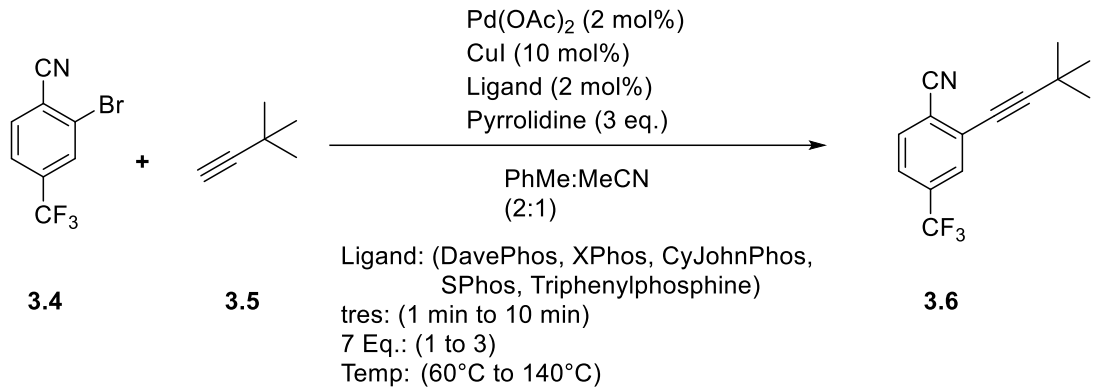


Scheme 8. Catalytic cycle for the Sonogashira cross-coupling reaction between an aryl or vinyl halide and terminal alkyne, facilitated by a palladium catalyst and copper (I) co-catalyst in a side catalytic cycle.

The Sonogashira cross-coupling reaction remains one of the most popular reactions for the formation of a sp^2-hybridised carbon atom with another sp-hybridised carbon atom. The utilisation of these produced arylalkynes holds significant importance in organic synthesis, particularly in the creating of natural products with various synthetic pathways involving the Sonogashira coupling. Moreover, the arylalkyne framework plays a critical role in the development of APIs within the pharmaceutical industry. Furthermore, the increased attention on the creation of aryl/vinyl alkynes has stimulated in-depth mechanistic studies aimed at enhancing the understanding of the reaction pathway and innovating novel catalytic approaches.[134,135] Utilising the Sonogashira cross-coupling to form these bonds offers inherent benefits, including its high tolerance of different reaction conditions and wide range of functional groups available. Consequently, the Sonogashira coupling reaction has emerged as an essential synthetic technique within the chemical industry.

Numerous factors can impact the performance of catalytic reactions, including sterics, solvent, ligand, heteroatoms, base, temperature and metal

source. This often necessitates significant optimisation for different substrate pairs. The Jensen group previously optimised Suzuki-Miyaura cross-coupling reactions performed on a droplet-flow microfluidic system utilising a mixed integer nonlinear programming (MINLP) optimisation approach. However, this method had limitations, focusing only on a single objective, and limiting overall process understanding.[87,102] Therefore, an efficient experimental approach for multi-objective optimisation of mixed variable catalytic systems on a substrate-by-substrate basis was sought after. Consequently, a pharmaceutically relevant Sonogashira cross-coupling reaction was chosen for investigation using the newly developed MVMOO self-optimisation approach.[136] This study focussed on the Sonogashira cross-coupling between aryl bromide **3.4** and terminal alkyne **3.5** to produce aryl alkyne **3.6** (Scheme 9).



Scheme 9. Sonogashira cross-coupling reaction between aryl bromide **3.4** and terminal alkyne **3.5** to form aryl alkyne **3.6**. Includes optimisation parameters.

Modifications were made to the original synthesis of a TRPV1 receptor antagonist, primarily used for pain management and treatment of chronic pain illnesses.[137–139] These adjustments were implemented to optimise the process using the automated mixed variable flow platform. These changes include: (i) substituting the aryl chloride with the corresponding aryl bromide to enhance the reaction rate, (ii) replacing the Pd_2dba_3 with the more stable $PdOAc_2$ precatalyst, and (iii) using a greener homogenous reaction mixture consisting of PhMe:MeCN (2:1) and pyrrolidine base instead of NEt_3. To ensure the stability of the metal-ligand complex, the prepared solution reservoirs were stored under nitrogen throughout the optimisation process.

3.3 Automated Flow Platform and Optimisation Results

The flow process operates through a custom written automated MATLAB script that has control over all of the equipment on the setup and the capabilities to monitor the performance of each piece. When adapting the setup for a new case study, it requires editing of the MATLAB optimisation script, new custom written conditions generation and response functions scripts in order to ensure the transfer of data to the algorithm is complete. However, with all the changes, the general operation of the automated script remains the same with separate timers throughout the optimisation which can act as indications where the script is erroring out if so. Throughout the duration of the analysis timer, the reagent pump flow rates are set to dead-time conditions to preserve material consumption. The switching valve is switched to a position containing only solvent to flush out any potential discrete variable remaining within the tubing line. Additionally, the solvent pump flow rate is set to 1 mL min^{-1} to ensure any residual components from previous iterations are flushed out. The addition of a 5-minute wait time once the pumps and switching valves had been set to their desired values ensured reagents would be at the desired concentrations and flushed through the system. The general operation of the script is shown below in Figure 31.

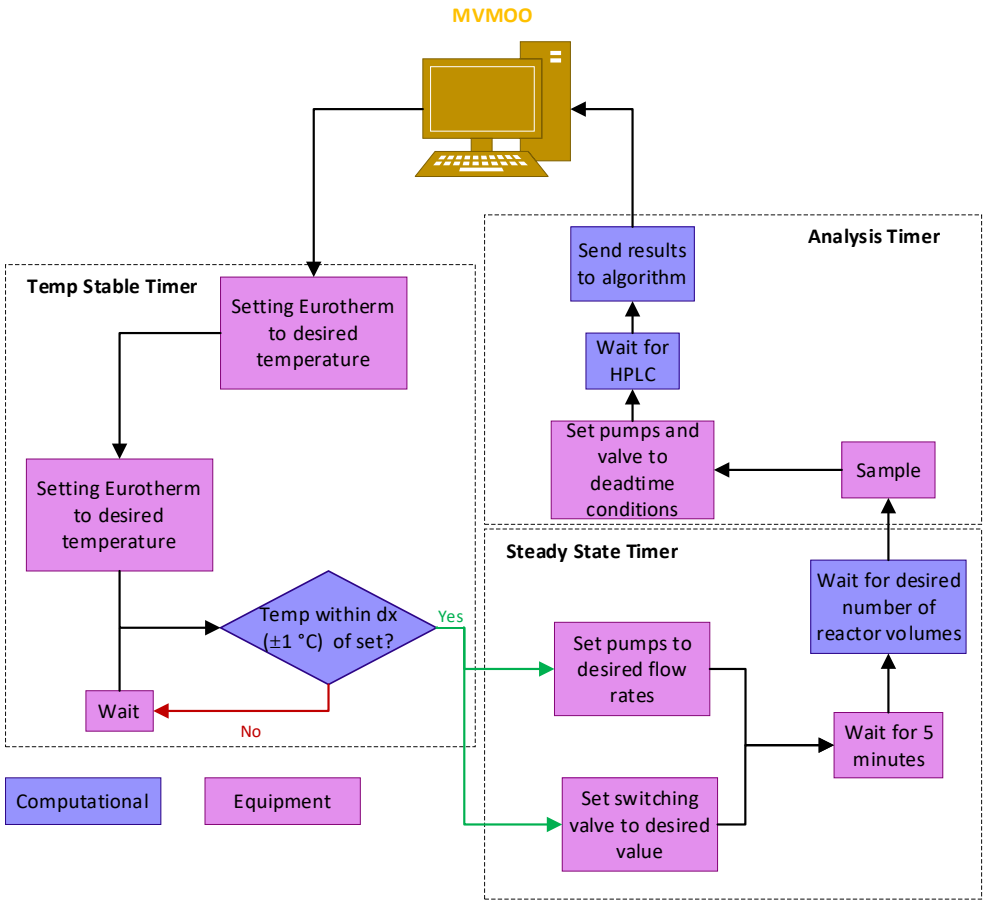


Figure 31. Flow chart for the MVMOO self-optimisation procedure with inclusion of details for the timers throughout the process.

Initial investigations in the optimisation began with determining a suitable set of process objectives to be optimised for, to do this a set of 12 initial experiments were completed on the flow platform. Using the data gathered from the discovery reactions, different process metrics could be calculated using the HPLC data, with a focus on finding a suitable economic metric. To determine whether they were a good fit, the metrics were all plotted against one another to ensure that a trade-off could be achieved, as shown in Figure 32. Initially, yield, cost and turn over frequency (TOF) were tested due to the interest in finding the best performing ligand for the associated cost of the reaction. This would have direct applications to the pharmaceutical industry where there is a critical desire to minimise associated costs per experiment and maximise the productivity of the reaction. To further explore the effects of the ligand on these initial studies, the ligand loading % was included to explore outcomes on the TOF and associated costs per experiment.

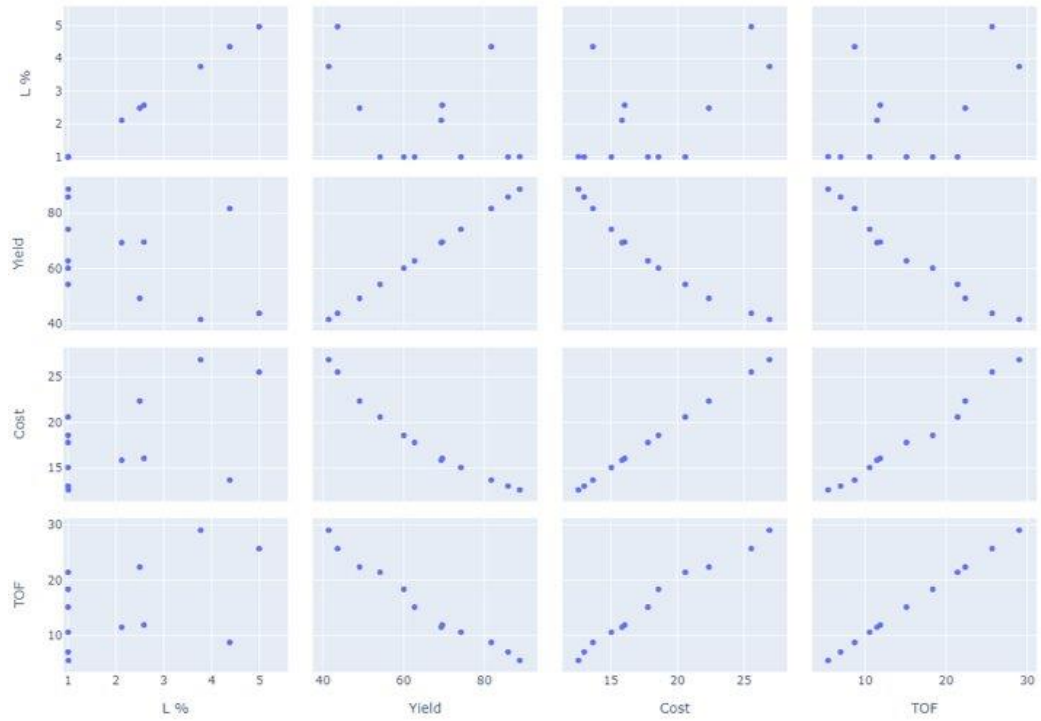


Figure 32. Exploration of different process metric outcomes using a 12 initial reaction data set, evaluating productivity metrics.

However, after later investigations, it was decided to include an industry relevant productivity metric and environmental process metric for the optimisation. Selection of an environmental metric was decided such that these processes would align and be applicable for enhancing the overall green aspect of these workflows. The productivity (STY) and an environmental (RME) metric both being selected as the objectives to identify viable operating conditions. RME is the percentage of actual mass of the desired product formed to the mass of all the reactants used, considering both atom economy and product yield.[140] Considering the pivotal role of ligands in catalysed reactions, the optimisation process incorporated the selection of a phosphine ligand as a discrete variable, alongside residence time, equivalents of the terminal alkyne (**3.5**) and temperature.

STY $$\frac{mass_{product}}{Volume \times t_{res}} \quad \textbf{(16)}$$

RME $$\frac{MW_{product\ \mathbf{3.6}} \times Yield}{MW_{\mathbf{3.4}} + (MW_{\mathbf{3.5}} \times equiv_{\mathbf{3.5}})} \quad \textbf{(17)}$$

$$minimise\ [-(RME), -(STY)] \quad \textbf{(18)}$$

Subject to:

$$Ligand \in [DavePhos, XPhos, CyJohnPhos, SPhos, TPP]$$

$$residence\ time\ \in [1.0, 10.0]$$

$$Equivalents\ (\mathbf{3.5}) \in [1, 3]$$

$$Temperature\ \in [60, 140]$$

To ensure comprehensive optimisation, a diverse set of monodentate phosphine ligands were chosen for investigation. These ligands were preferred due to their excellent activity in palladium-catalysed cross-coupling reactions, good solubility in organic solvents, and easy accessibility in the commercial market. Moreover, ligands with varying cone angles were included in the selection to investigate the impact of steric bulk on the reaction's outcome. Additionally, ligands with identical bonding stoichiometry to the catalytic centre were selected to ensure a valid and comprehensive comparison between them was achieved. The palladium catalyst chosen was related to this cone angle, where sterically bulky ligands have the capacity to induce the dissociation of the active palladium catalyst from the inactive resting state. As a result, these ligands have the potential to facilitate the formation of palladium complexes from palladium species that are weakly coordinated, such as $Pd(OAc)_2$ or $Pd_2(dba)_3$.

When comparing the two potential Pd catalysts, $Pd_2(dba)_3$ can be characterised by the sterically bulky *trans*,*trans*-dibenzylidene acetone (dba) ligands, which provides steric protection to the Pd centre. This steric protection restricts the potential for ligand exchange, as the presence of these bulky dba ligands restricts the approach, thereby reducing the exchange rates.[141] Additionally, the dba ligands are electron-rich, resulting in strong coordination to the Pd centre and enhancing the stability of the complex, further reducing the rate of ligand exchange compared to its acetate counterpart. Both combined can lead to the requirement for long stirring times and heating required to achieve full ligand exchange, but in addition, the non-innocence of the dba ligand can lead to the presence of $Pd(dba_x)L_y$ type complexes, which is highly undesirable for this optimisation.[142] In contrast,

$Pd(OAc)_2$ features acetate ligands that are often more susceptible to displacement by other ligands. Consequently, presenting $Pd(OAc)_2$ as the preferred catalyst for these investigations, where full ligand exchange is crucial to form the desired Pd-L_2 discrete variable complexes. The inclusion of triphenylphosphine as selection for the final ligand provided a cheaper alternative ligand of choice when compared to the remaining more costly Buchwald phosphines. By deciding this, it would allow for comparison between the associated cost and performance of each ligand within the optimisation. This would further provide useful insight for the pharmaceutical industry where price per experiment vs. outcome performances are crucial for economic analysis. Additionally, to the varying steric bulk, the Buckwald phosphines were chosen to include ligands with a range of electron density.

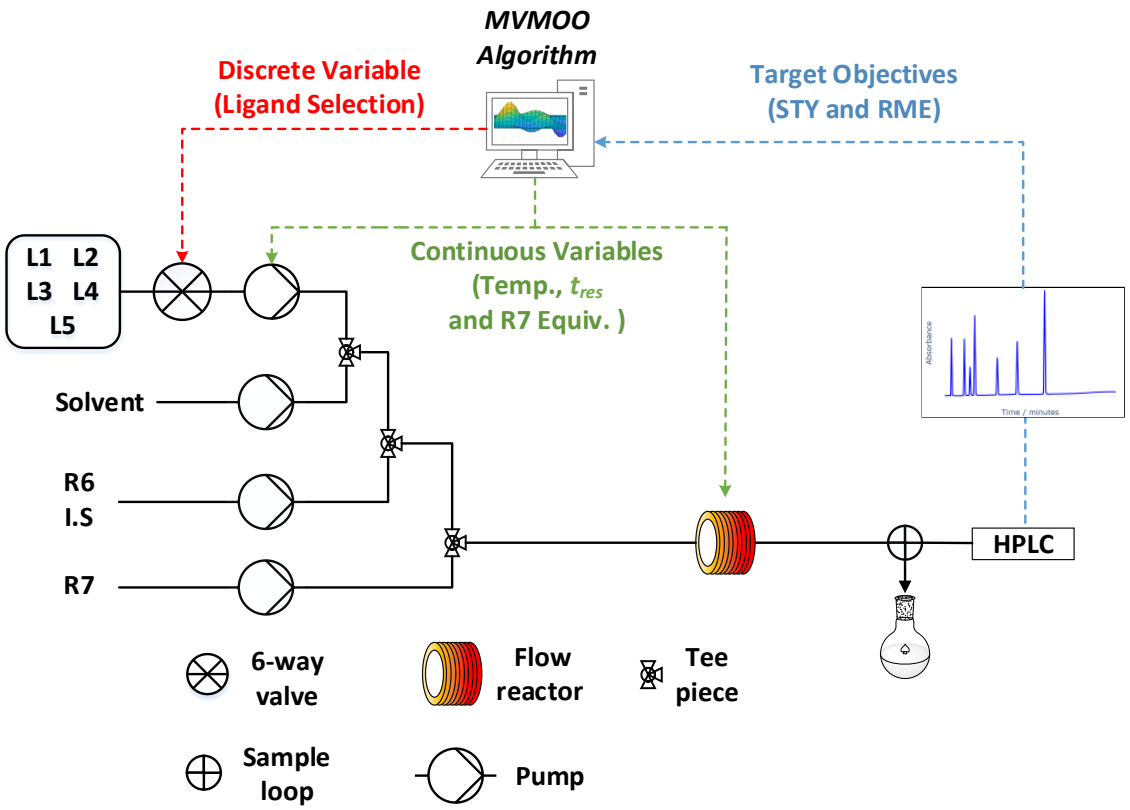


Figure 33. Schematic used for the optimisation of the Sonogashira cross-coupling case study. Where L1 is DavePhos, L2 is XPhos, L3 is CyJohnPhos, L4 is SPhos and L5 is TPP. The catalyst is $Pd(OAc)_2$, CuI and pyrrolodine base were included in each of these stock solutions. Additionally, R6 represents 2-bromo-4-(trifluoromethyl) benzonitrile (**3.4**), I.S is the internal standard 1,3,5-trifluoromethoxybenzene and R7 represents 3,3-dimethylbutyne (**3.5**).

The MVMOO algorithm was initialised with 25 experiments, employing five LHC experiments per ligand, and then running a further 44 sequential iterations. Among these experiments, 12 were identified as non-dominated solutions, illustrating the trade-off between STY and RME as shown in Figure 34. The optimal STY achieved was 322.0 kg m^{-3} h^{-1}, accompanying a corresponding RME value of 51.5. On the other hand, the optimal RME obtained was 68.2, with a corresponding STY of 32.31 kg m^{-3} h^{-1}. In this case, the Pareto front reveals a distinct and steep linear trade-off, enabling a significant improvement in STY and only a minor detrimental effect on RME. Remarkably, the optimum RME conditions achieved an outstanding in-situ yield of 90%, as shown in Figure 35. This highlights the applicability of

optimising green metrics such as RME within these processes, where the optimum RME relates to less waste produced within the reaction due to higher efficiency, whilst still producing excellent yields within the system.

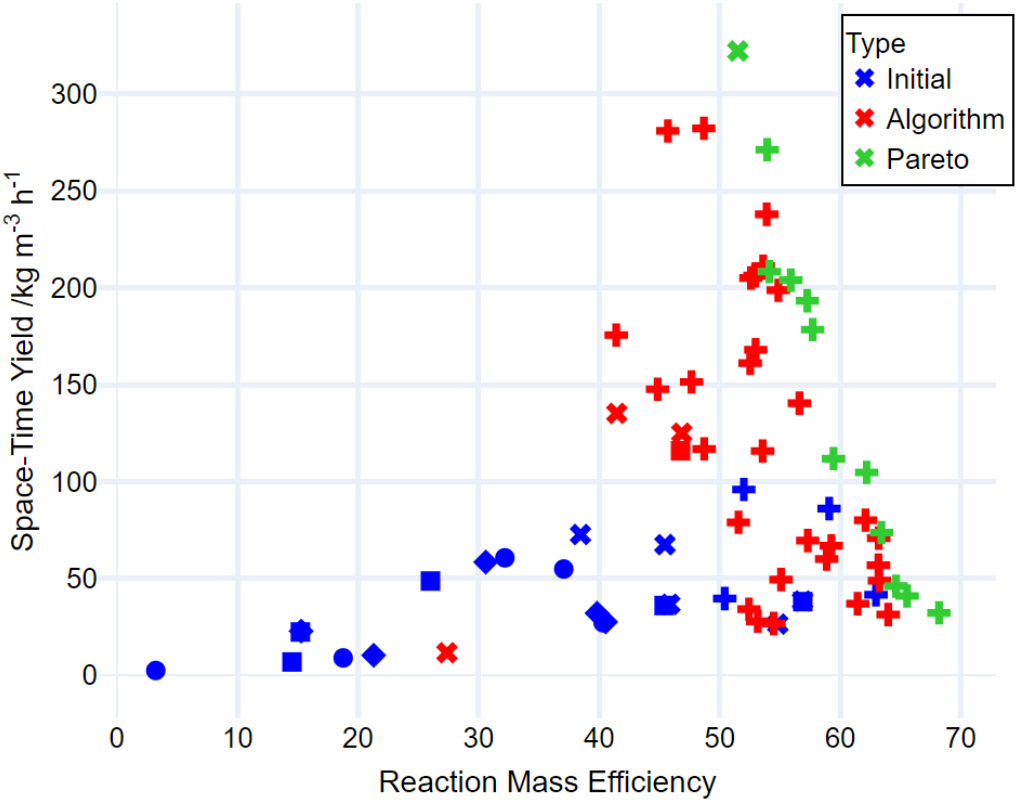


Figure 34. Results of the four-parameter mixed variable multi-objective optimisation of the Sonogashira cross-coupling reaction. An initial 25 experiments were completed from LHC sampling with an additional 44 iterations direct by MVMOO, 12 of which formed a Pareto front highlighting the trade-off in STY and RME. Ligand shapes represent: ● – DavePhos, ◆ – XPhos, ■ – CyJohnPhos, × – SPhos, + – TPP.

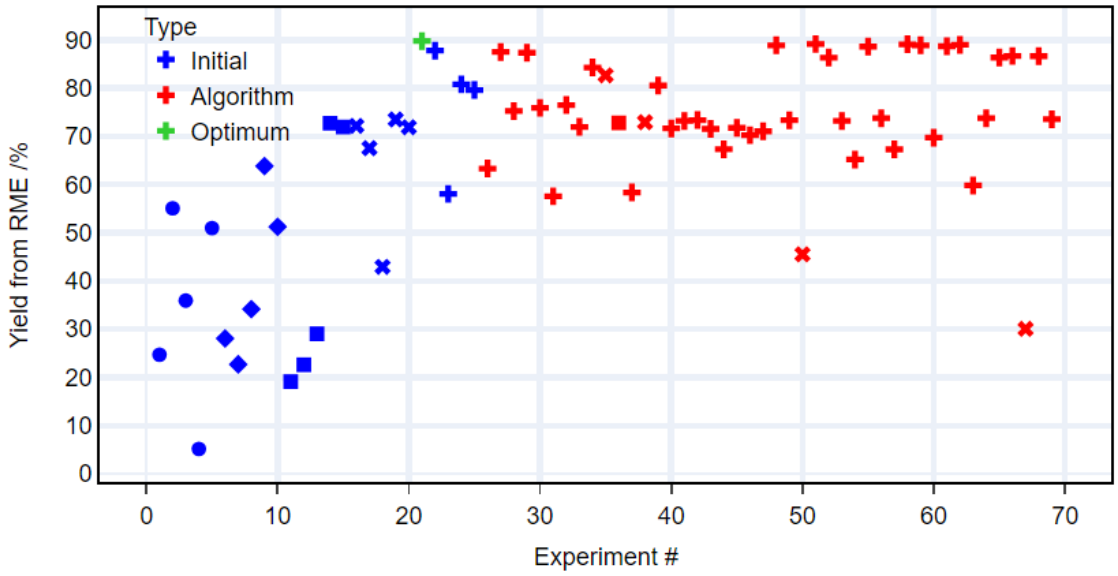


Figure 35. Plot of the calculated yield from RME values vs experiment number, highlighting the outstanding yield achieved of 90% using the conditions corresponding to the optimum RME using TPP. Ligand shapes represent: ● – DavePhos, ◆ – XPhos, ■ – CyJohnPhos, × – SPhos, + – TPP.

The initial LHC results fell within the objective space region characterised by a range of RME with low STYs. Notably, triphenylphosphine consistently outperformed all other ligands in terms of both STY and RME in each LHC experiment. As a result of this, the MVMOO algorithm predominantly recommended triphenylphosphine-based experiments for subsequent iterations. Generally, Sonogashira cross coupling reactions involving deactivated aryl halides favour the use of sterically demanding and electron-rich phosphine ligands. The cone angle serves as a measure of the steric bulk of ligands, and the monodentate phosphine ligands used in this study followed the trend: TPP < DavePhos, XPhos < CyJohnPhos < SPhos according to literature values. Therefore, the optimisation results challenge conventional chemical understanding, demonstrating the importance of efficient experimental optimisation, especially in complex reactions involving novel substrate pairs where interactions are not fully understood or easily predictable.

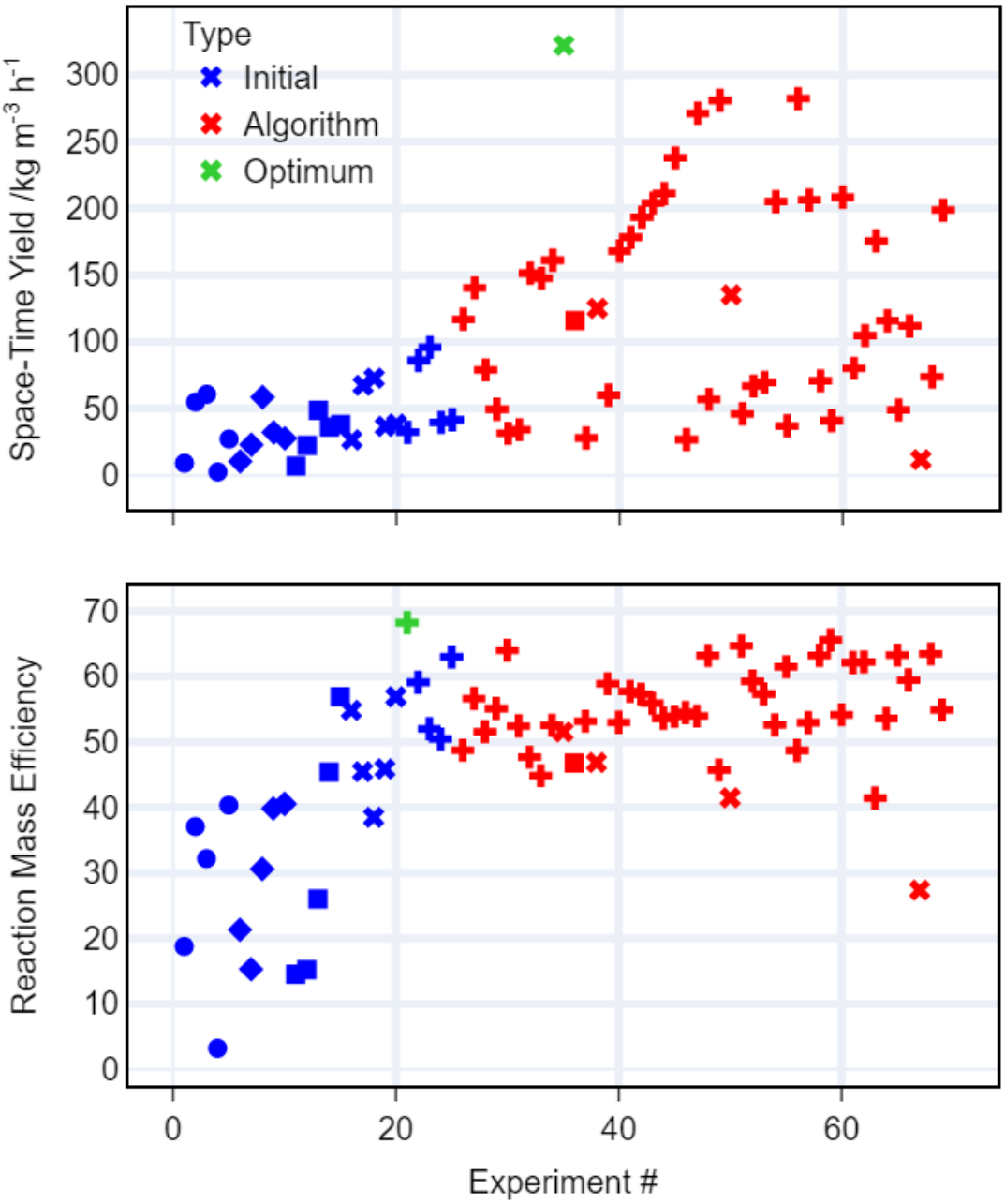


Figure 36. Results from the Sonogashira coupling optimisation with individual process metrics against experiment number to highlight metrics enhancements over the optimisation process. Where ■ – Initial, ■ – Algorithm and ■ – Optimum.

Despite the initial LHC focussing on regions of low STY, the MVMOO algorithm successfully identified new regions in the objective space corresponding to more productive regions which may have been missed by other experimental techniques, Figure 36. This underlines the efficacy of the MVMOO approach for exploring and discovering better reaction conditions to

optimise process metrics. The achievement of the optimum RME point within the initial space filling LHC, facilitated the MVMOO algorithm to improve on the STYs for the primary iterations. These improvements were highlighted by the identification of the optimum STY within the first 10 iterations increasing initial highest performing STY of 95 kg m^{-3} h^{-1} up to 322 kg m^{-3} h^{-1}. Upon identification of the optimum STY, the algorithm then effectively explored the trade-off between the objectives, subsequently mapping out a further 10 Pareto front points within just 34 experiments. The late iterations within the optimisation highlight the reproducibility to achieve high RME results. The MVMOO algorithm showcased its impressive exploratory capabilities by conducting a limited number of experiments using SPhos as the ligand, which was identified as the second most promising ligand in the system during the initial LHC. Notably, the highest achievable STY during the optimisation was achieved when utilising SPhos at low residence times, high equivalents, and high temperatures, Figure 37.

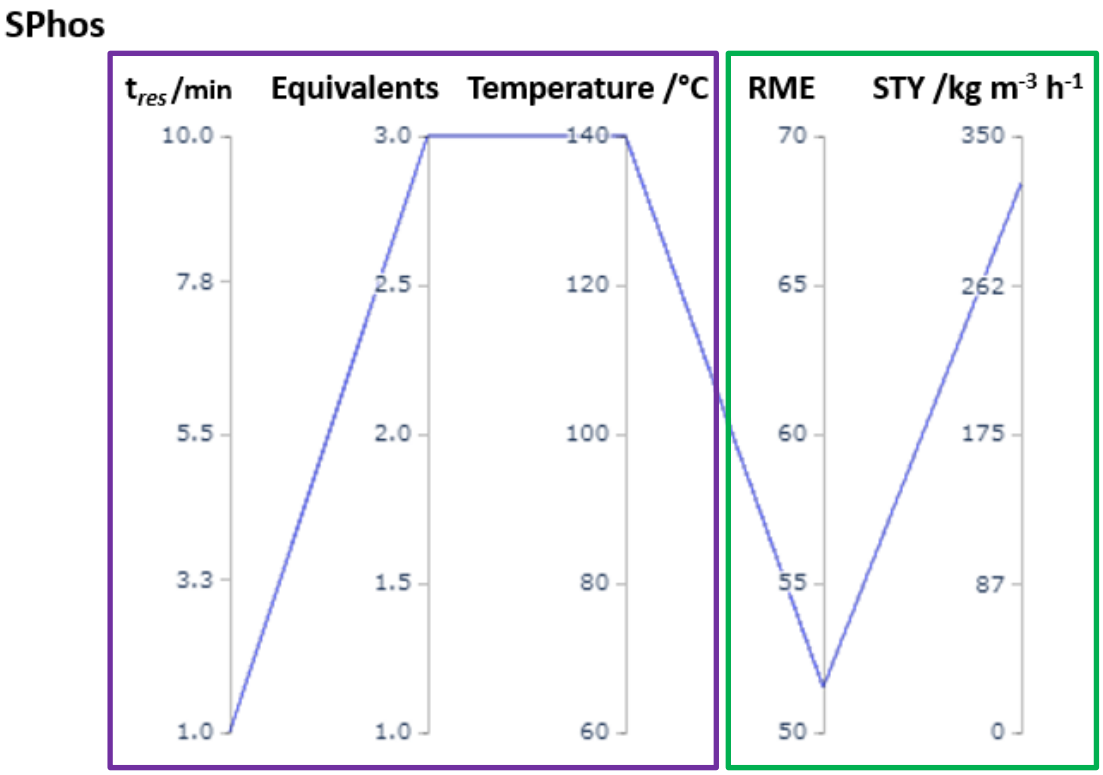


Figure 37. Parallel coordinates plot showing the interactions between the variables for the Pareto optimal solution for the STY that was achieved by SPhos.

Subsequently, the MVMOO algorithm proceeded to directly compare these continuous conditions utilised on SPhos while employing triphenylphosphine, which resulted in a less favourable STY and RME as a consequence of lower yield. In contrast, triphenylphosphine exhibited a preference for employing low to moderate equivalents of alkyne (**3.5**), combined with high residence times and low temperatures to achieve higher RMEs. Alternatively, it favoured shorter residence times and higher temperatures to maximise STYs, and in contrast to the high SPhos related STY, TPP utilised a range of low-moderate equivalents to achieve this, Figure 38.

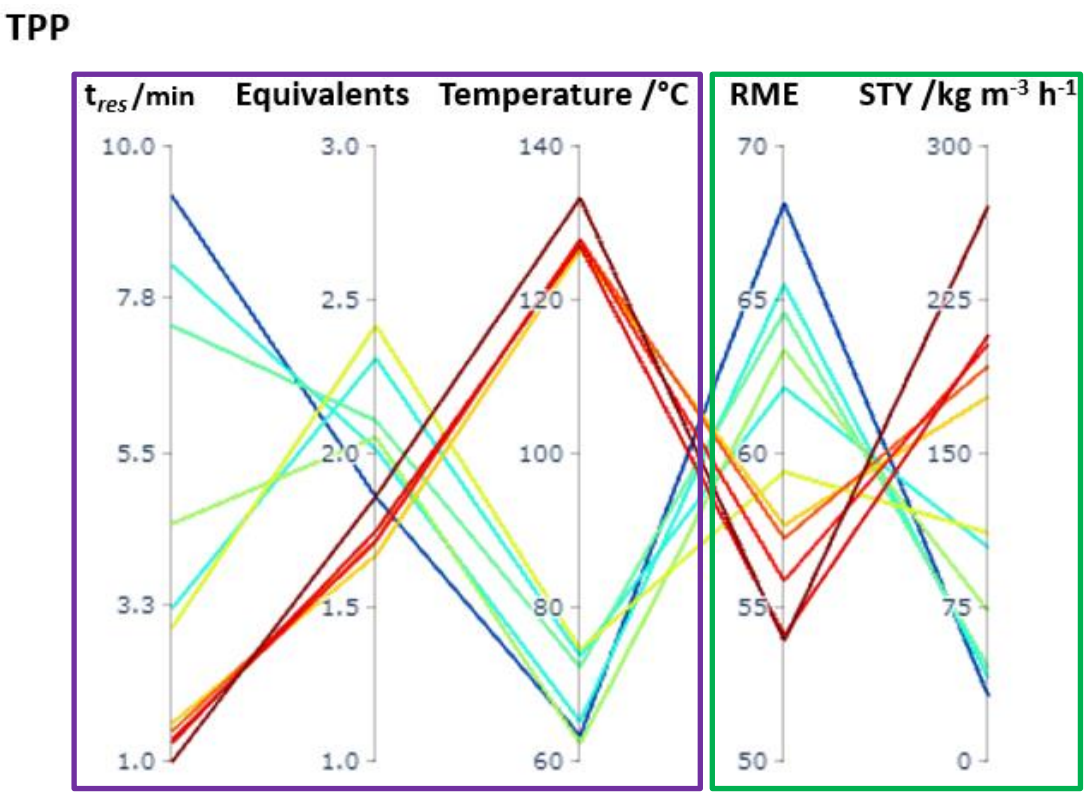


Figure 38. Parallel coordinate plot highlighting the interactions between the variables for each non-dominated solution of the Sonogashira optimisation. Each line represents a single Pareto optimal solution for TPP ligand points. Line colour is scaled in relation to STY weighting to aid in visualisation (high STY/low RME = –, low STY/high RME = –).

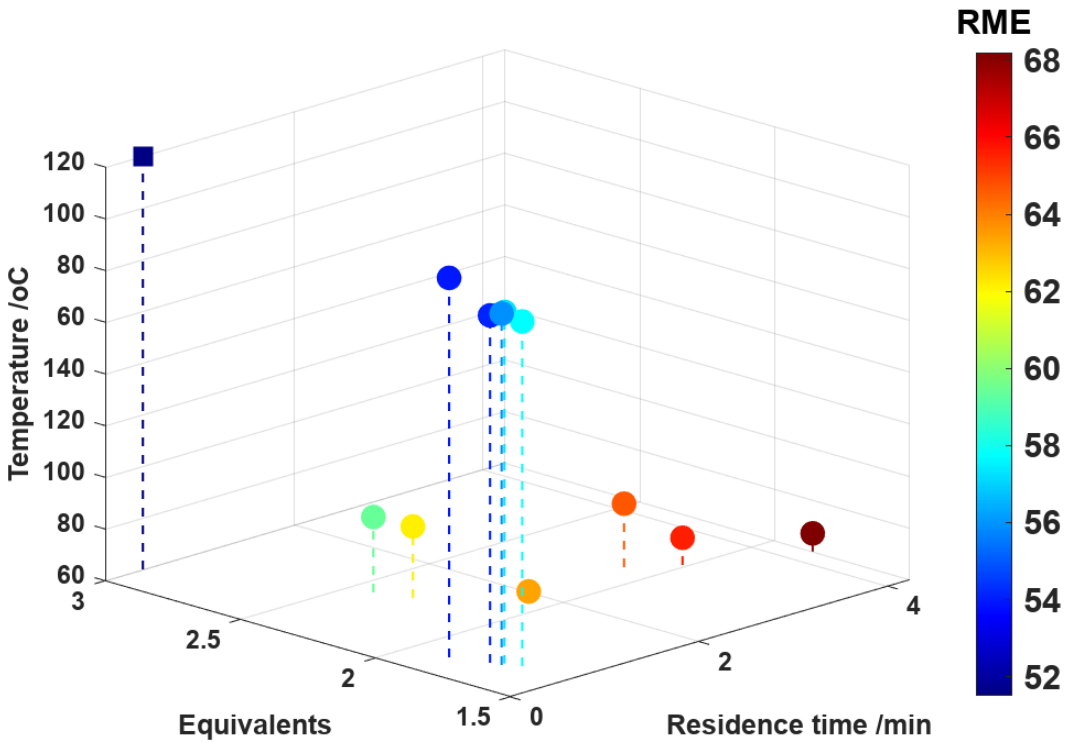


Figure 39. 3-D plot for the Pareto front solutions with respect to RME, where ● represents TPP based solutions and ■ represents the singular SPhos solution.

Similar to the S_NAr case study (Chapter 2), it became evident that the interactions between the continuous variables and each ligand were significantly distinct, highlighting simultaneous optimisation of mixed variables as a superior approach for identifying the true optima compared to the conventional sequential optimisation method. This capability to fine-tune complex mixed-variable catalytic reactions for diverse substrate pairs not only opens up exciting prospects for exploring new regions of chemical space and novel reactions but also remains viable without the requirement for excessively labour-intensive experimentation in the manor which traditional sequential optimisations inherently require. In this study presented, the MVMOO self-optimisation methodology that was developed exhibited the capability to autonomously optimise a four parameter, Sonogashira cross-coupling reaction, including consideration of both discrete and continuous variables. This optimisation was achieved with respect to two objectives with remarkable efficiency, requiring only 69 experiments conducted over the span of 22 hours. Notably, there was no human intervention required beyond the initial preparation of reagent stock solutions.

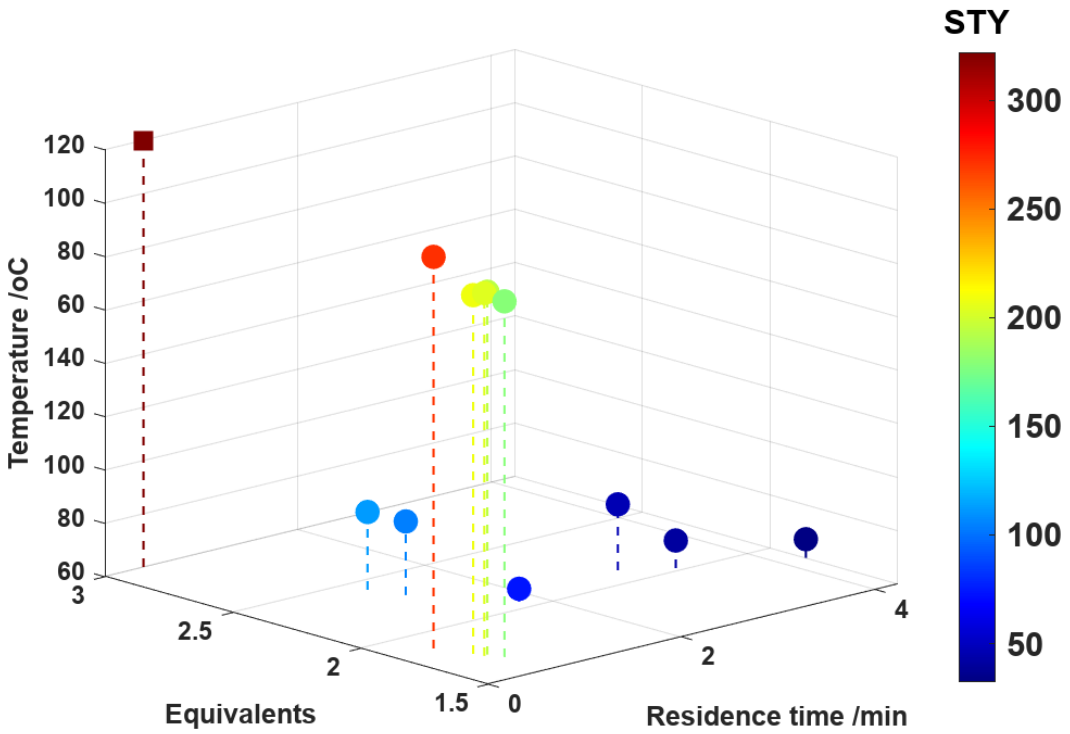


Figure 40. 3-D plot for the Pareto front solutions with respect to STY, where ● represents TPP based solutions and ■ represents the singular SPhos solution.

Incorporation of real-time monitoring of the hypervolume progression after 60 experiments allowed for the optimisation to be terminated once a relative plateau of 5 experiments was reached, Figure 41. Implementing this monitoring and predefined stopping criterion ensured a balanced trade-off between the information gained in the optimisation and minimisation of time and material consumption. This trade-off is particularly crucial for catalytic optimisations with high associated costs linked to expensive catalysts and ligands. Performing the optimisation with the criteria in place, facilitated the identification of a further 3 Pareto optimal solutions after the 60th iteration leading to the increase in hypervolume, and information gained that can observed in this region.

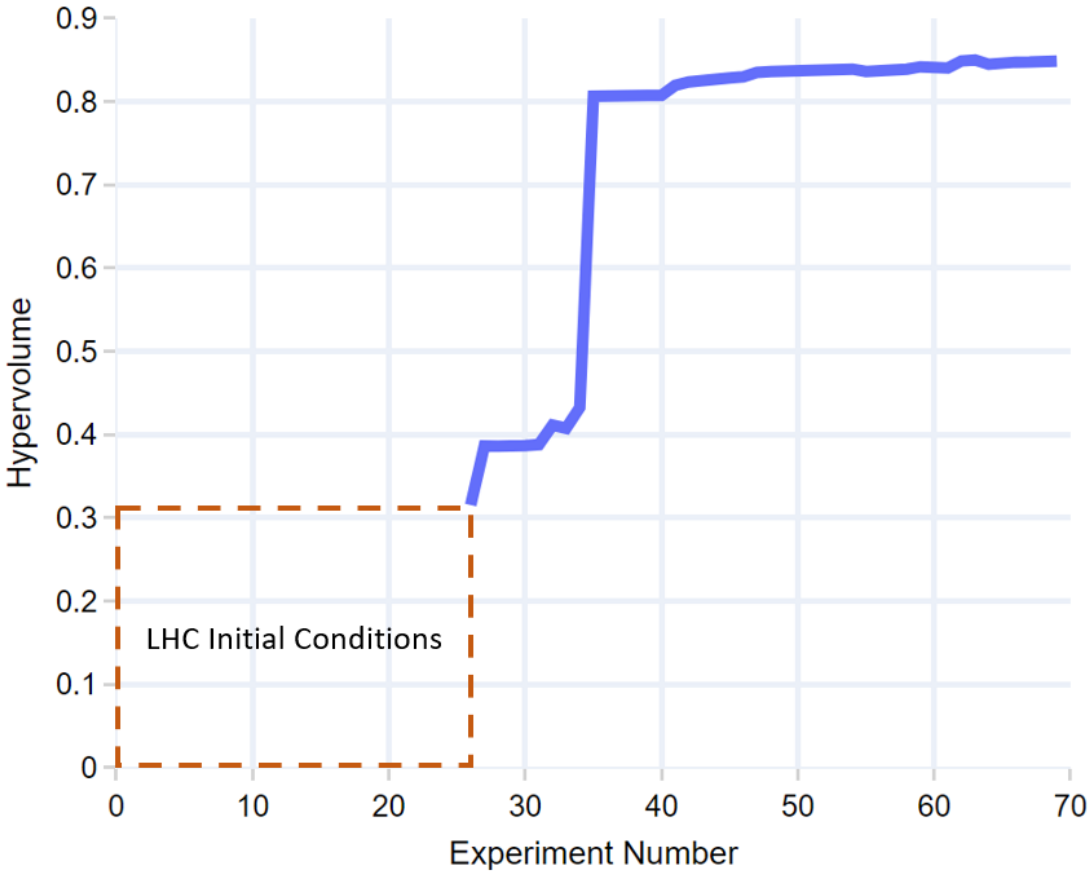


Figure 41. Calculated hypervolume vs experiment number for the Sonogashira optimisation that was used to determine the stopping criteria for this workflow.

The efficiency and effectiveness demonstrated in this autonomous optimisation workflow is comparable to previously reported experimental optimisations of single objective mixed variable systems and multi-objective continuous variable studies.[66,87] Therefore, it is anticipated that this type of approach for complex catalytic optimisations will prove effective in expanding the range of available tools for synthetic and process chemistry alike.

3.4 Conclusion

The utilisation of self-optimisation algorithms for the optimisation of mixed-variable catalytic systems has been previously developed on several varying metal catalysed reactions.[74,79,87] However, these studies solely focussed on single-objective approaches. This type of workflow eliminates the potential to provide understanding into the trade-off between conflicting performance metric, whether environmental or productivity. Insight into this

trade-off between objectives is crucial for the development of feasible industrial processes, which aim to balance different performance criteria. Consequently, this leaves the necessity to develop multi-objective mixed-variable optimisation approaches for the consideration of catalytic systems. In this work, an automated multi-objective optimisation was explored for the Sonogashira cross-coupling reaction, with the inclusion of continuous and discrete variables in the system. Ligand selection as the discrete variable was based on the crucial role they play in catalysed reactions. Various monodentate phosphine ligands with a range of cone angles were selected owing to their prominent activity in palladium catalysed reactions in addition to good organic solvent solubility and commercial accessibility.

The optimisation aimed to explore and maximise a productivity metric (STY) and environmental objective (RME) to gain insight into the trade-off between the conflicting criteria. Inclusion of both types, aligned with the evolving pharmaceutical industry interests, in which the design of environmentally greener synthesis is desired without significant reduction to the productivity of the reaction. A pharmaceutically relevant Sonogashira cross-coupling reaction was selected, where the product was an intermediate for the synthesis of TRPV1 receptor antagonists, used for the management and treatment of pain. The MVMOO algorithm successfully explored the Pareto front for the conflicting objectives in 69 experiments, of which 12 were identified as non-dominated solutions, over a 22-hour period. An optimum STY of 322.0 kg m^{-3} h^{-1} was achieved utilising SPhos in conjunction with low residence times, high temperatures and equivalents of alkyne. Conversely, an optimum RME of 68.2 was achieved using TPP alongside high residence times, low temperatures, and low-moderate equivalents of alkyne. Remarkably, this optimum RME result corresponded to a 90% yield, highlighting the feasibility of achieving high RME whilst maintaining high production yields. Despite localisation of experiments on low STYs during the initial LHC, the MVMOO algorithm effectively identified new regions in the objective space, representing more productive regions. Identification of TPP as the overall best performing ligand for maximising the RME and exploring trade-off, except in the case of optimum STY, challenges conventional chemical understanding where the least sterically demanding ligand was

identified to have the best performance. Therefore, demonstrating the importance of efficient multi-objective experimental optimisation, where in a single-objective optimisation of STY, these observations would not have been identified. The significance of these multi-objective methodologies is especially true in cases featuring complex reactions including novel substrate pairs, where the interactions are not fully comprehended or readily predictable.

Chapter 4. Enhancing Telescoped Chemical Reactions through Mixed Variable Optimisation

4.1 Introduction

In the ever-ongoing objective for more streamlined and environmentally friendly chemical processes, the field of synthesis has undergone a notable transformation, leading to the emergence of sequential synthesis in the chemical industry.[143] This emergence has been heightened by the integration of continuous flow processes in the wider chemical industry which boosts efficiency over single-pot synthesis and allows access to the efficient synthesis of new chemical scaffolds.[12,144] Performing sequential synthesis, enables the simplification of chemical synthesis, a reduction in waste and time required by eliminating individual work-up and purification steps between reactions, providing suitable analytical techniques have been developed and optimised for these processes. Traditional chemical synthesis requires discrete separate steps for all reactions involved in the full synthesis, with each step requiring individual optimisation and purification processes.[145] Mitigating these steps and combining them into a single process, bypasses the extended time requirements, heightened energy consumption and substantial waste generation.[118] Telescoped synthesis challenges these processes by performing numerous sequential reactions over a continuous process, increasing the overall efficiency, reducing the time requirements, and mitigating isolation, and purification steps. Elimination of these steps significantly reduces the solvent requirements, where solvents for the production of pharmaceuticals are estimated to account for 50% of greenhouse gas emissions.[118,146]

The pharmaceutical industry can significantly benefit from the inclusion of telescoped synthesis where it can be observed that it can boost the speed and reduce the cost at which complex drug molecules can be synthesised and developed.[147] For the development of complex molecules, sequential synthesis can aid in the assembly of functional groups and chemical scaffolds in single reaction steps without sacrificing the purity or efficiency of those steps. Compared to other chemical industries, the pharmaceutical manufacturing of APIs has historically displayed a higher level of environmental impact, highlighting a necessity for the adoption of alternative greener techniques.[148]

Large complex molecules in the pharmaceutical industry also benefit from the sequential techniques via combining protection or deprotection of functional groups that can occur in single reactions where further functionalisation can happen downstream, reducing waste and time requirements. These types of processes can accelerate drug discovery timelines while reducing the dependency on hazardous reagents or solvents to further assist in enhancing the efficiency, sustainability, and safety of the synthesis.[21,149,150] Presenting these types of processes as a crucial tool for expediting drug discovery and averting potential supply chain disruptions inherent in the traditional iterative batch methodology.

One interesting reaction that has been performed using telescoped methodology was by Baxendale et al. in which they were able to develop the *in-situ* generation of ethyl isocyanate which is typically a highly light and moisture-sensitive compound.[151] This *in-situ* generated compound was later reacted with different functionalised diazonium compounds via a subsequent cycloaddition for the synthesis of 1,2,4-triazoles. In addition, this methodology was further expanded onto forming pyrrolo[1,2-*c*]pyrimidines scaffolds which have pharmaceutical significance for the treatment of central nervous system disorders.

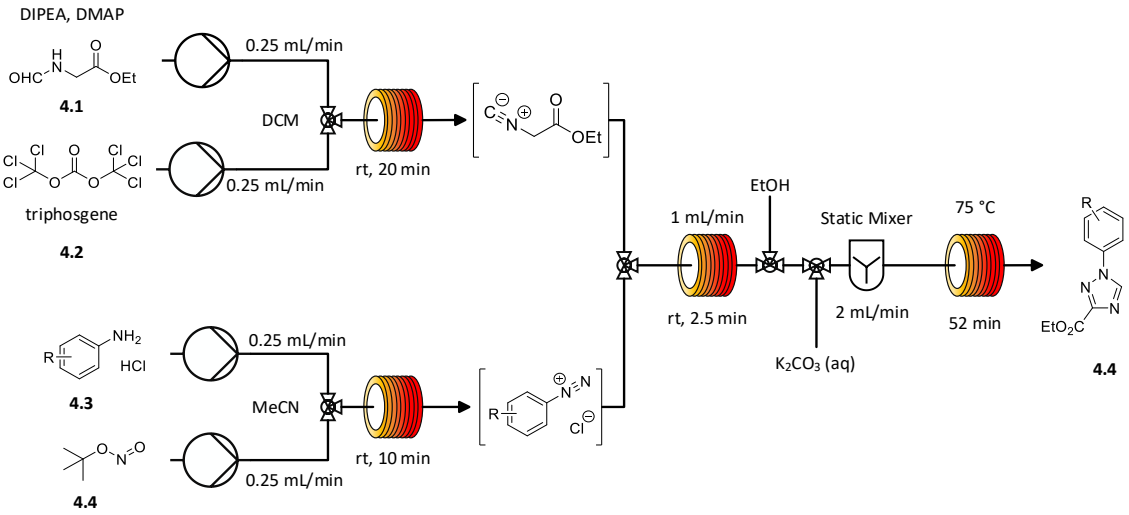


Figure 42. Flow schematic for the synthesis of 1,2,4-triazoles using continuous telescoped flow techniques.[151]

This work highlighted the effectiveness of utilising continuous telescoped flow techniques for the formation and downstream utilisation of challenging

components such as ethyl isocyanates. Whereby further reacting downstream, the authors were able to minimise any potential degradation of this sensitive species and any exposure of researchers to its potentially harmful effects, providing a safer synthesis for the desired compounds. However, this work presents some of the challenging tasks for sequential synthesis in which the optimisation of the reactions can lead to resource and labour-intensive studies, although it still precedes traditional batch methodologies. Optimising these types of reactions remains highly challenging due to the complexity of issues that the system presents, concatenating steps into a single process increases the number of reaction variables present which can negatively affect downstream processes, overall reducing yields or other metrics for each reaction step.[152] Additionally, the presence of prior chemicals or intermediates from previous steps can pose a threat to the downstream reactions via complex interactions with reagents that can further impact the efficacy of the current reaction step. Therefore, this suggests that for such complex multistep processes, efficient optimisation cannot be achieved by independently optimising reaction conditions. Instead, a comprehensive optimisation of the entire process, considering all reaction variables, is necessary.[153]

Approaching telescoped optimisations with the use of algorithms holds the potential to efficiently synthesise complex natural products over multiple steps whilst being able to control and optimise each step for the desired outcome. Using this self-optimisation approach with algorithms enables the process to consider any complex interactions present, identify them and provide an autonomous technique to rapidly develop telescoped reactions.[20] The rising potential to unlock efficient optimisations for these telescoped systems, has led to efforts being made to attempt to employ algorithms to automate these processes. The initial source of success was made by employing a sole analytical measurement of the product stream. Employing this approach, enabled the global optimum to be identified with the best operating conditions for the telescoped synthesis. However, by only sampling the final outlet stream, it critically restricts the process understanding gained for the specific stages occurring within the entire procedure. It limits any identification of crucial intermediates formed within any stage of the prior

reactions and limits the ability to understand the continuous variable's influence on the process metrics being studied. Advancements have been effectively made by Jensen *et al.* and Kappe *et al.* to demonstrate the ability to apply multiple inline and online analytic techniques to these types of optimisations.[152,154] Although success was made during these optimisations, using two separate analytical techniques leads to increased equipment costs which can substantially limit the application of these technologies to the wider research and industry community. Furthermore, employing inline analytical instruments can lead to extensive prior characterisation and validation requirements to achieve accurate measurements.

A recent publication by Clayton *et al.* has successfully demonstrated the ability to apply multipoint sampling to telescoped optimisations, overcoming the previously initially presented limitations.[20] This technique employed two sampling valves to monitor both stages of a two-step Heck cyclisation-deprotection reaction. Utilising a single piece of online analytical equipment, surmounts the issues previously relating to increased costs and prior workloads. To achieve this multipoint sampling, each reactor outlet was fed into individual sampling valves which were daisy-chained together to feed each reaction step samples into an HPLC instrument for analysis on a combined chromatogram, Figure 43.

Although this enabled both steps of the reaction process to be monitored and analysed for key intermediates and determination of process metrics over both stages, the process understanding gained is still limited due to the constraints of only considering continuous variables. Eliminating the potential of exploring discrete variables, critically limits the understanding of the key interactions between the mixed variables, which in telescoped optimisations can provide significant insights into the reaction pathways and mechanisms as well as highlight important intermediate formations. The incorporation of mixed variables into these algorithm-based telescoped optimisations is an area which has yet to be explored but arguably holds vast potential to revolutionise these types of optimisation processes. By integrating both discrete and continuous variables within the optimisation frameworks, a more comprehensive design space can be addressed. This approach not only enhances the efficiency and flexibility of the optimisation process but also

enables the exploration of the relationships and interactions between the variables over the whole telescoped reaction pathway. As such, there remains a necessity to develop an efficient methodology for multi-objective optimisation within telescoped flow reactions which include discrete variables alongside continuous parameters within the optimisation domain.

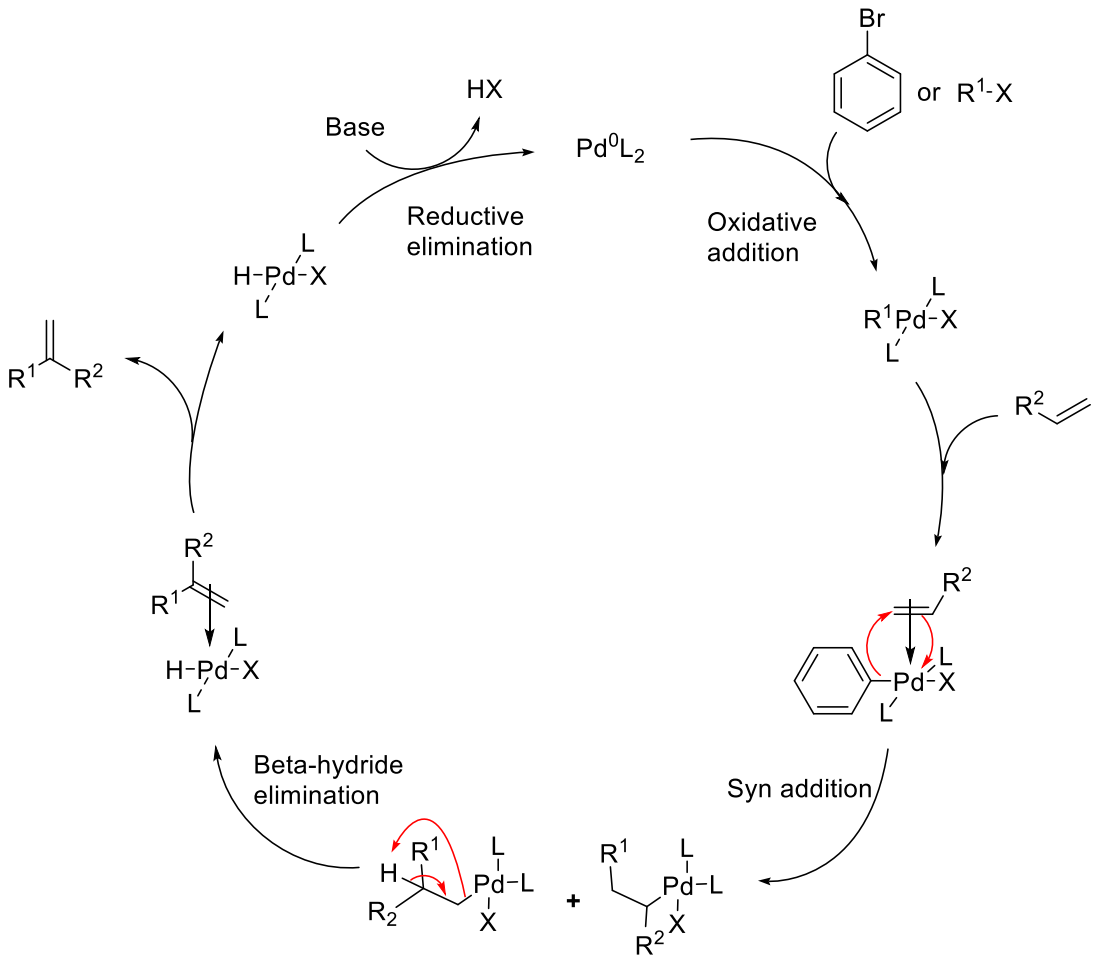
This chapter advances the research initiated in the preceding chapters by tackling the presented challenges posed for telescoped optimisations. This is achieved through the exploration of mixed-variable optimisations for a telescoped flow reaction. The choice of the case study was deliberate, aiming to extend the literature work completed by Clayton *et al*. on the Heck-Hydrolysis telescoped reaction and enhance it via the introduction of a mixed-variable system.[20] Expanding on the results from Chapter 3, which emphasised the importance of the ligand effects on catalytic systems, the decision to persist with the selection of ligands as the discrete variable was made. To comprehensively account for the interactions between variables across both steps of the telescoped reaction, the objective metrics were carefully chosen. These metrics encompass the productivity of the initial step and the overall yield of the final product resulting from the combined steps.

4.2 Heck Reaction

The Heck reaction, also known as the Mizoroki-Heck reaction stands as one of the most critical advancements made in organic synthesis reactions, first developed in late the 1960s and early 1970s by Richard Heck and Tsutomu Mizoroki.[155] This was highly important as it pioneered palladium-catalysed coupling reactions for the construction of carbon-carbon bonds. It was later recognised as such an important development in organic synthesis that it was awarded the Nobel Prize in chemistry in 2010 alongside the Suzuki-coupling reaction. This innovative reaction involves the cross-coupling between aryl or vinyl halides with alkenes which react in the presence of a palladium catalyst and base which has led to the synthesis of a diverse range of substituted alkenes over years of implementation.

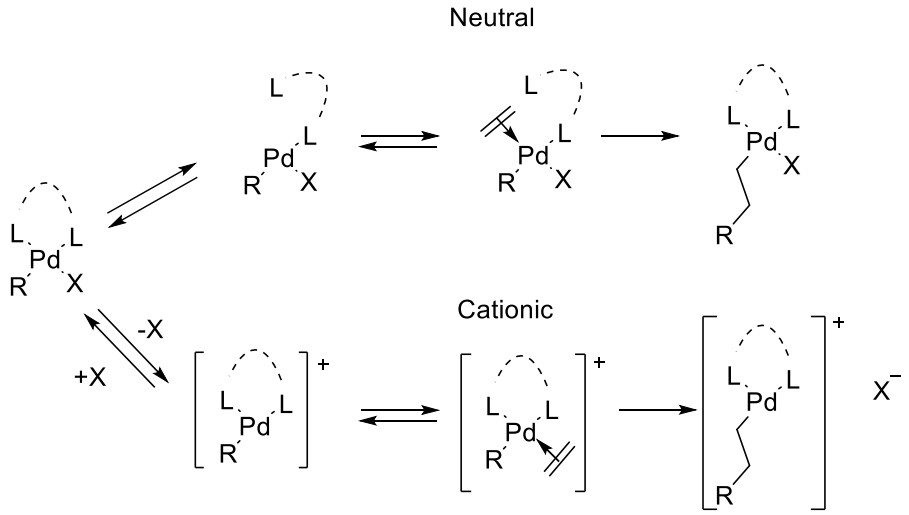
The reaction often requires a pre-activation step of a Pd(II) catalyst pre-cursor before the catalytic cycle can begin, this is initiated by the reduction of

the Pd(II) complex to the active Pd(0) species. Upon formation of the active Pd(0) complex, the reaction is initiated through an oxidative addition of the palladium complex to the aryl or vinyl halide. Generally, the presence of electron-donating phosphine ligands can aid in the activation of the Pd(0) catalyst to help accelerate the rate of oxidative addition, this can be further increased by the selection of the halide with general reaction rates observed being: I >OTf >Br>Cl.[156]



Scheme 10. Catalytic cycle for the Heck reaction between an aryl or vinyl halide and alkene, facilitated by a palladium (0) catalyst which is regenerated via the addition of a base in the final step of the cycle.

Following this, the cycle undergoes the coordination of the alkene to the palladium complex followed by a syn addition of the alkene, which requires dissociation of the ligands. This mechanism is notably influenced by the selection of the phosphine ligands and halides incorporated in the reaction path. The pathway branches into two routes: one being a cationic and the other a neutral route. While monodentate ligands can lead to both pathways, bidentate ligands tend to favour the cationic route. However, the neutral mechanism for bidentate ligands can still occur providing the ligand of choice has a large bite angle.[157]The insertion of the alkene into the palladium centre is an important step in the catalytic cycle. Its significance stems from its power to control the stereo and regio-selective outcome of the reaction, ultimately shaping the final product. Regioselectivity is heavily influenced by the steric bulk on each side of the alkene during the insertion.



Scheme 11. Neutral and Cationic pathways for alkene insertion within the Heck catalytic cycle for the direction of alpha and beta products.[157]

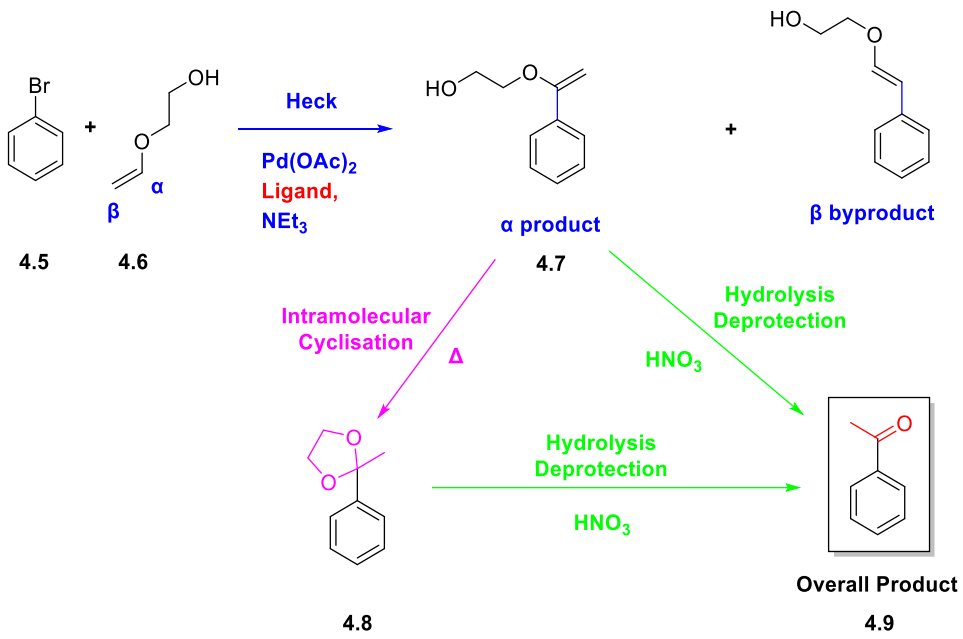
In general, the least sterically hindered side of the alkene results in the formation of the bond with the palladium complex. This is the general case for neutral palladium complexes, however, when the palladium complex is cationic, the major product formed is influenced by a trade-off of the steric and the electronics of the alkene, with nucleophilic attack occurring on the side of the least electron density. Following the insertion of the alkene into the complex, a beta-hydride elimination takes place providing that the hydride is attached syn-coplanar to the palladium centre. Upon elimination, it yields the new substituted alkene, with the E-configuration favoured due to unfavourable

steric interaction in the transition state for the Z-isomer. In the final step of the cycle, the addition of the base is required to cause the reductive elimination step, where the palladium centre is reduced back to its $Pd(0)L_2$ state and the regeneration of the active catalyst is complete.

Over the years, the Heck reaction has been extensively researched and pivotal in complex natural synthesis, the formation of APIs and materials innovation.[158,159] Its exceptional versatility and capacity to facilitate various functional groups have enabled the reaction to become an essential instrument for synthetic chemists. This heightened attention has led to the reaction being involved in several optimisation processes aimed at enhancing its efficiency and expanding the scope of the reaction over different processes. The area of automated optimisation has embraced this reaction extensively, which can be observed by numerous literature publications intended at pushing the boundaries of optimisation and exploiting the Heck reaction for various process metrics and continuous parameters available.[40,50,145,160,161]

Selection of this reaction was made as the design of this telescoped case study was to follow on and advance the work completed by Clayton *et al.* on their two-step telescoped Heck-cyclisation-deprotection reaction.[20] As previously mentioned, the exclusion of discrete variables in this pathway limits the understanding of the process and the key interactions that could be derived from the initial optimisation procedure. Therefore, leading to opportunities for further exploration, particularly when considering mixed variable systems with respect to multiple objectives. The pathway for this reaction began with an initial Pd-catalysed Heck reaction yielding the intended regioselective product. Subsequently, an intermolecular cyclisation took place to generate a ketal, followed by the second step of the telescoped reaction, where the focus was to hydrolyse the ketal using nitric as a catalyst to form the desired final product. Modifications were made to the starting material of choice to use bromobenzene in place of the aryl bromide originally used, this stemmed from constraints due to its availability and cost of the original aryl bromide. In the original proposed reaction scheme, the approach aims at synthesising a potential precursor for 1-methyltetrahydroisoquinoline C5 functionalised derivatives which have pharmaceutical interest for the treatment of depression.[162] Furthermore, nitric acid was substituted for the

original tosylic acid. This substitution was also influenced by the lack of a *t*-boc protecting group in the new starting material, where nitric acid's capability to deprotect this group led to an alternative acid being used in the original synthesis. Consequently, given the absence of the Boc group in the new compound, the utilisation of nitric acid for the hydrolysis reaction was selected. Although this modified approach changes the original starting material, this study is aimed at highlighting the possibility and potential of optimising mixed variable cases on these types of systems, where a change of the initial compound used could result in the unlocking of these types of optimisations for potential API pathways. The fully modified two-step reaction pathway is shown in Scheme 12.



Scheme 12. Heck reaction between bromobenzene (**4.5**) and ethylene glycol vinyl ether (**4.6**) to form the alpha substituted product (**4.7**) followed by intramolecular cyclisation to form the dioxolane product (**4.8**). Subsequently, the hydrolysis deprotection reaction to form the final acetophenone product (**4.9**).

4.3 Telescoped Optimisation

For this telescoped optimisation, the automated flow platform that had been used for the previous case studies saw the addition of a second unheated reactor where the second step hydrolysis deprotection reaction

would take place. Additionally, a second sampling valve was added to allow access to analyse both steps of the reactions, these were oriented in a daisy-chained orientation, similar to the configuration that Clayton *et al.* had implemented, Figure 43.[20] This enabled access to both steps of the reaction, which would allow process metrics to be calculated and optimised for either step of the reaction pathway. To facilitate the use of a single HPLC instrument for the comprehensive analysis of the two-step reaction, a dual-method approach was employed. Initially, upon sampling from the first valve, a 7.3-minute gradient method was executed transitioning from 16.3% MeCN to 95% MeCN (83.7% water to 5% water) to elute the target compounds effectively. Following the completion of the first HPLC method, the system was reset to the initial analysis conditions, preparing for the subsequent injection. Subsequently, the delayed second valve sampled and executed a second HPLC method of the same gradient conditions, extending the analysis for an additional 7 minutes.

By employing this sequential dual-method approach, a consolidated HPLC chromatogram was generated for the entire two-step reaction. This combined chromatogram was then processed using the optimisation script, enabling the identification and quantification of the desired signals corresponding to the reaction products and intermediates. This streamlined analytical workflow facilitated accurate process metric calculation and optimisation, leveraging the capabilities of a single HPLC instrument for comprehensive analysis and data interpretation.

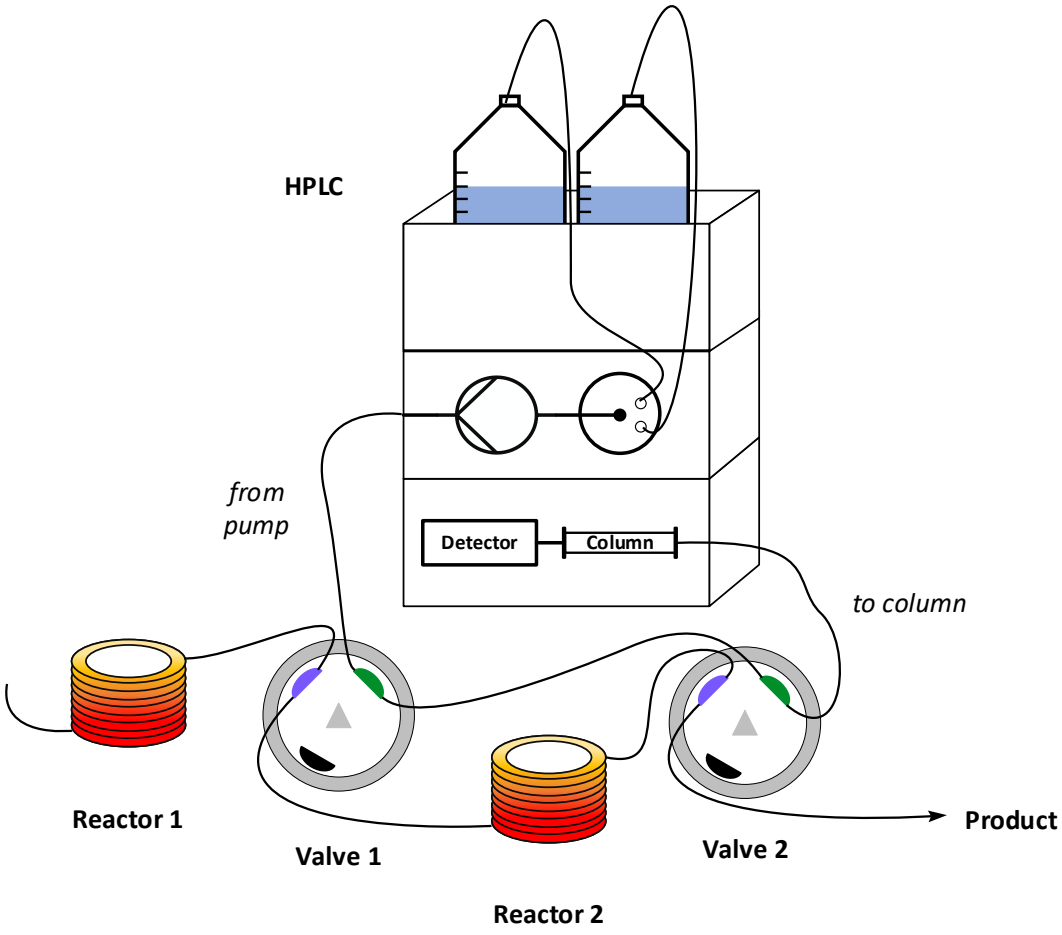


Figure 43. Multipoint sampling technique used for telescoped optimisation with daisy-chained Vici sampling configuration where a single HPLC analysed both steps in the reaction. Configuration based on the work completed by Clayton *et al.*[20]

The process metrics for the optimisation were selected so that a metric from either step of the reaction was chosen to provide a unique optimisation study, to evaluate the interactions over the whole process. The overall yield of the acetophenone (**4.9**) and STY for the formation of the dioxolane from the Heck reaction (STY_{Diox}) were selected, with the latter being selected due to the Heck reaction being the rate-limiting step of the process. For the STY_{Diox} process metric, the decision was made to utilise the dioxolane as the input instead of the alpha intermediate. Although both undergo hydrolysis to yield acetophenone (**4.9**), the intermediate is more readily hydrolysed which would restrict the continuous variable limits, especially the temperature. This choice was made to ensure a comprehensive exploration of the temperature range set in the optimisation, thereby enabling the observation of its effect on both objective metrics. Optimising this step, in addition to overall yield, highlights

the viability of these types of workflows to the pharmaceutical industry, in which the desire to maximise product formation and minimise time requirements is of high importance. Furthermore, this approach would assist in further exploring the trade-off between the objectives, where increased cyclisation of the intermediate to enhance higher STY_{Diox} would, in turn, lead to slower rates of hydrolysis for the overall product formation. This could result in a fascinating objective space, requiring the MVMOO algorithm to effectively exploit the continuous variable interactions to explore suitable trade-offs of both objectives.

The useful insights gained from the Sonogashira coupling optimisation that had been completed in Chapter 3, led to the decision for the ligand to be selected as the discrete variable of choice for manipulation in this telescoped case study. In this workflow, the ligand selection approach would deviate from that of the Sonogashira optimisation. This stems from the potential to generate both an alpha- and beta-product via the Heck reaction. In this context, the selection of the ligand would further highlight the necessary interactions between the discrete and continuous variables for the formation of the desired alpha intermediate (**4.7**) and dioxolane (**4.8**). This would have a direct effect on the process metrics, in which both are dependent on the alpha intermediate and dioxolane formations within the reaction pathway.

To increase the complexity of the reaction pathway and overall process, it was decided to include a mix of monodentate and bidentate ligands for selection, with a range of diphenylphosphine-based ligands selected due to the impressive performance of 1,3-bis(diphenylphosphino)propane (dppp) in the continuous variable Heck-hydrolysis optimisation performed by Clayton *et al.*[20] In addition to dppp, bis(diphenylphosphino)methane (dppm) and 1,2-bis(diphenylphosphino)ethane (dppe) was selected as the other two bidentate ligands to explore how the increase in ligand size would affect the outcome of the reaction in conjunction with poor solubility in the solvent choice observed for other larger variations. For the monodentate ligands, TPP was selected due to its performance in the Sonogashira optimisation and DavePhos was selected as it has shown efficiency in promoting the Heck reaction with aryl bromides.[163] To capture the effects of the continuous variables over the whole process, it was decided to incorporate parameters from both the initial Heck

reaction and the subsequent hydrolysis step. These included the residence time for the whole process, equivalents of ethylene glycol vinyl ether (**4.6**), the temperature within the first reactor, and the ratio between the acid flow rate and the flow rate of the first reactor, Table 5.

Table 5. Boundaries for continuous variables in the telescoped Heck-hydrolysis optimisation.

	Residence time (min)	*Equivalents of vinyl ether (**4.6**)*	*FvA : FvR1*	*Temperature (°C)*
Lower bounds	1	1	0.5	120
Upper bounds	20	3	1.5	200

Due to the rapid nature of the hydrolysis step both substituents, it was decided not to include any temperature parameters for this step as ambient reactor temperature was sufficient. This led to the following optimisation equation and variables.

$$minimise\ [-(STY_{Diox}\ (\mathbf{4.8})), -(Overall\ Yield\ (\mathbf{4.9}))] \quad \textbf{(19)}$$

Subject to:

$$Ligand \in [dppm, dppe, dppp, DavePhos, TPP]$$

$$residence\ time\ \in [1.0, 20.0]$$

$$Equivalents\ (\mathbf{4.6}) \in [1, 3]$$

$$Ratio\ \in [0.5, 1.5]$$

$$Temperature\ \in [120, 200]$$

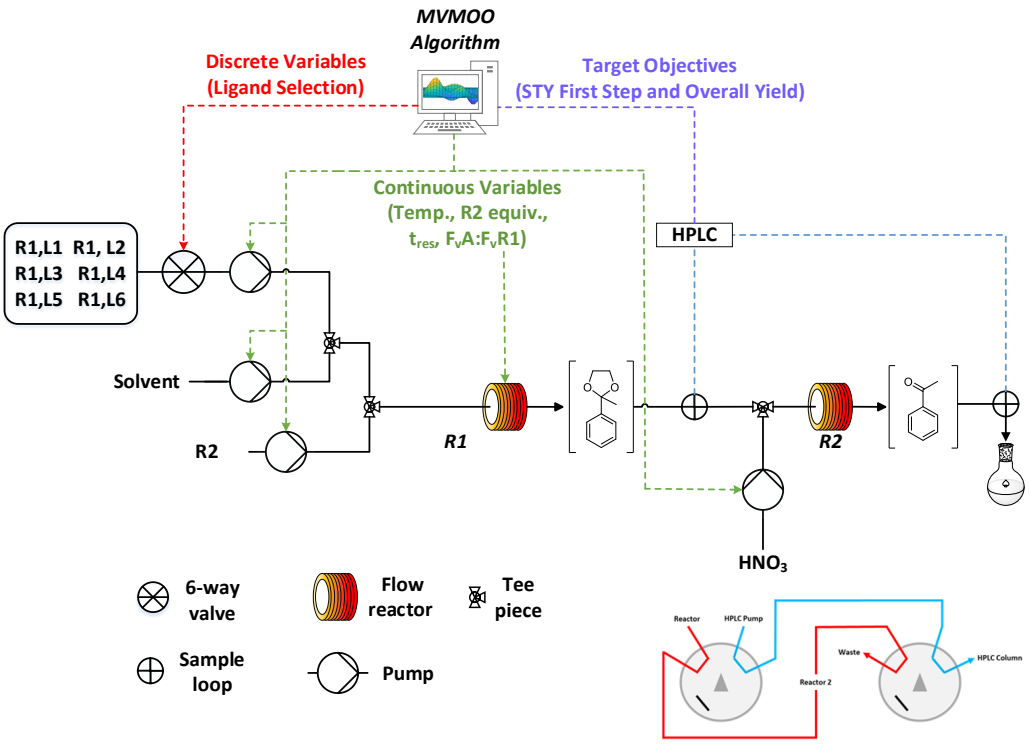


Figure 44. Flow schematic used for the telescoped optimisation case study. Where R8 is bromobenzene (**4.5**), L1 is dppm, L2 is dppe, L3 is dppp, L4 is DavePhos and L5 is TPP. The catalyst $Pd(OAc)_2$, NEt_3 base and internal standard methyl p-tolyl sulfone were included in each of these discrete variable stock solutions. Additionally, R9 represents ethylene glycol vinyl ether (**4.6**), and the solvent is ethylene glycol:MeCN (2:1 ratio).

The MVMOO algorithm was initialised with 25 experiments, employing five LHC experiments per ligand in the same format employed in Chapter 3. The algorithm then sequentially ran through a further 41 experiment iterations achieving the identification of 8 non-dominated solutions sufficiently highlighting the trade-off between the STY_{Diox} and overall yield, Figure 45.

In this case, it can be observed that the Pareto front reveals a steep non-linear trade-off between the objectives, where major improvements to the STY_{Diox} acts significantly detrimental to the overall yield formed. Where the overall yield is the yield of the acetophenone (**4.9**) formed from the multi-step reaction, through the hydrolysis deprotection reaction of either the α-intermediate (**4.7**) or the subsequent dioxolane product (**4.8**). The conflicting objective metric effects are emphasised by the fact that all experiments yielding more than 70% overall, resulted in a STY_{Diox} less than 5. Conversely,

this is further highlighted by the observation that every experiment with an STY_{Diox} exceeding 15 kg m^{-3} h^{1} attained an overall yield in the range of 28% to 43%, emphasising the negative effect each objective had on the other. For this study, dppp presented itself as the dominant ligand of choice, achieving both optimal results for STY_{Diox} and overall yield and governing the Pareto front, with all non-dominated solutions associated with this ligand. The optimal STY_{Diox} achieved was 40.5 kg m^{-3} h^{-1} which corresponded to an overall yield of 32%. Conversely, the optimal acetophenone (**4.9**) overall yield achieved was an impressive 88% yield which related to a poor STY_{Diox} (**4.8**) value of 2.97 kg m^{-3} h^{-1}. The initial LHC results spread across a region of space characterised by a range of moderate overall yields and low STYs except for two high-performing STYs achieved by dppe and dppp.

As previously mentioned, dppp consistently outperformed all other ligands in the system in terms of STY_{Diox} (**4.8**) and overall yield with dppe being the only other ligand to closely contest it. This resulted in the MVMOO algorithm mainly deciding for dppp to be the ligand of choice for subsequent iterations after the LHC was complete. In this telescoped reaction, the hydrolysis reaction requires the formation of the alpha intermediate (**4.7**) during the Heck reaction. This alpha intermediate (**4.7**) can then be hydrolysed to yield acetophenone product (**4.9**). Alternatively, it can undergo cyclisation to generate the dioxolane (**4.8**) product which can also be hydrolysed, with the intermediate being more readily hydrolysed compared to the ketal. This cyclisation transformation occurs within the synaddition step of the cycle, where the regioselective coordination of the alkene determines this formation.

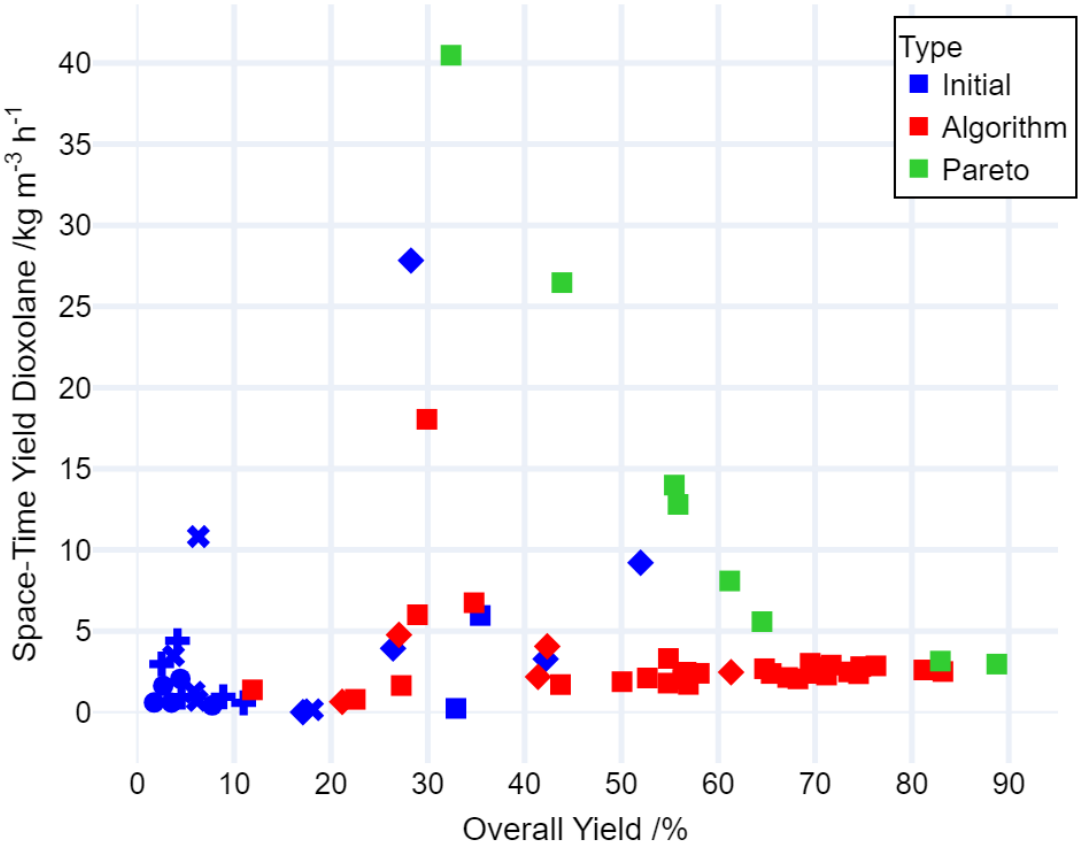
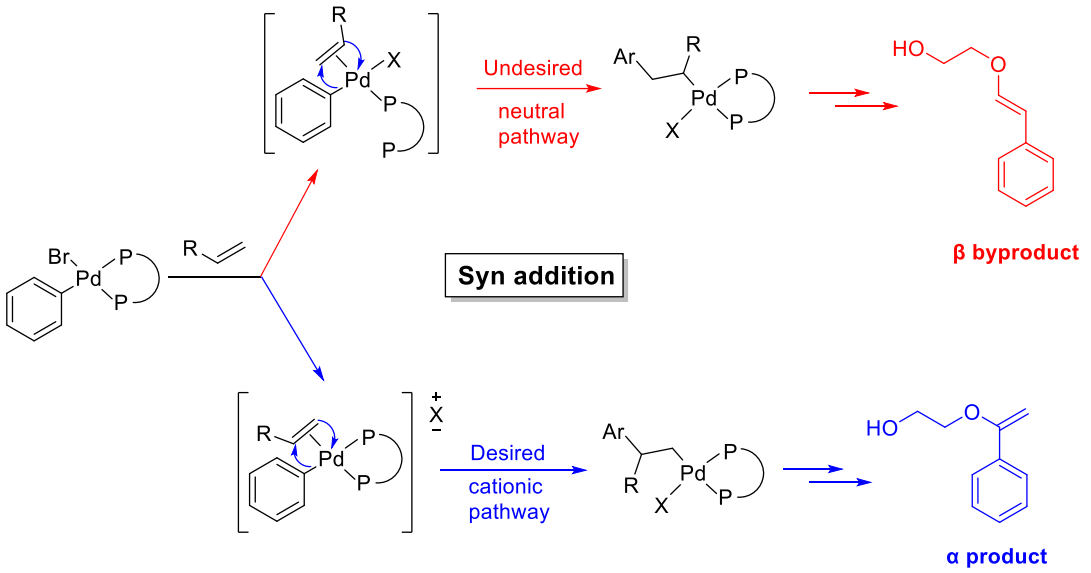


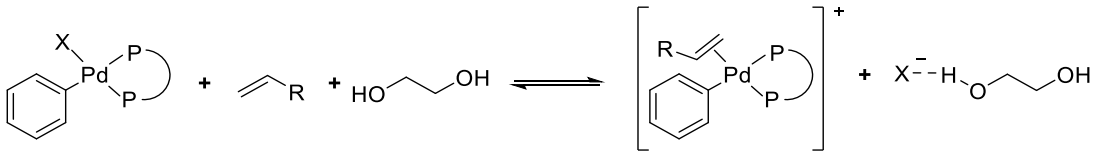
Figure 45. Results of the five-parameter mixed variable multi-objective optimisation of the telescoped Heck-hydrolysis reaction. For each ligand, 5 initial LHC points were afforded. The MVMOO algorithm then iteratively ran an additional 41 experiments, 8 of which were identified as non-dominated solutions highlighting the trade-off between STY_{Diox} (**4.8**) and acetophenone (**4.9**) overall yield. Ligand shapes represent: ● – dppm, ◆ – dppe, ■ – dppp, ✕ – DavePhos, + – TPP.

To accomplish this generation of the alpha intermediate, the beta position of the alkene must coordinate with the palladium centre. This leads to the already coordinated phenyl group being connected to the alpha position of the alkene, such that upon beta hydride elimination, the desired alpha intermediate is formed from the cycle. Typically, this type of coordination is achieved through the cationic pathway which is facilitated by the dissociation of the halide from the palladium centre. This leads to a nucleophilic attack on the least electron-dense side of the alkene by the cationic palladium centre, which in this case is the alpha position.



Scheme 13. Formation of the desired alpha- and undesired beta-products via the respective cationic and neutral pathways in the Heck reaction.

A proposed literature concept suggests the presence of an equilibrium stage preceding the rate-determining step, which is likely the alkene insertion.[164] Within this equilibrium, the introduction of a potential hydrogen-bonding donor is believed to shift the balance of the equilibrium in favour of the cationic palladium(II)-alkene intermediate.[165] As a result, the concentration of this cationic intermediate increases, thereby promoting a more rapid generation of the alpha product. This relates to this system where the presence of ethylene glycol as part of the solvent composition can enable hydrogen bonding within this equilibrium to promote alpha formation.



Scheme 14. Equilibrium for the formation of the cationic palladium(II)-alkene intermediate promoted by the presence of hydrogen-bond donation by ethylene glycol.

As previously mentioned, the desired cationic pathway is typically facilitated by bidentate ligands, which could rationale the subpar performance exhibited by the monodentate DavePhos and TPP throughout the optimisation. Furthermore, in the case of bidentate ligands, the bite angle serves as a measure of their steric bulk, similar to that of the cone angle for monodentate ligands. Among the bidentate ligands utilised in this

optimisation, the following trend for bite angle emerges according to literature values: dppp < dppe < dppm.[166] This bite angle directly impacts the steric crowding around the palladium centre, further promoting the alkene insertion to take place at the alpha site. Consequently, this elucidates why dppp stands out as a prominent ligand throughout this optimisation process. Additionally, this rationalises the outcomes noted for dppe; although it surpasses its less sterically bulk counterpart dppm, whilst still falling behind the effectiveness of dppp.

The MVMOO algorithm's successful identification of dppp as the optimal ligand within the LHC emphasises the effectiveness of telescoped optimisations, where minimisation of material consumption is crucial. Demonstrating the ability to manage and manipulate complex interactions and help guide the formation of key intermediates throughout multiple-step procedures, particularly when these intermediates and interactions are not straightforward to predict due to the inherent complexity of telescoped processes. This emphasises the proficiency of these types of approaches in controlling the interactions between the factors at different stages, allowing for the effective exploration and discovery of optimal reaction conditions to optimise the process metrics set within the telescoped pathway. Although the LHC returned relatively low results for the overall yields, the identification of two high STY_{Diox} results in the initial dataset, enabled the algorithm to explore regions of high overall yields to maximise the hypervolume of the objective space. By doing this, the algorithm was able to successfully identify new regions in the objective space relating to higher overall yields with significant improvements made to this metric, where the maximum overall yield of 61% from the LHC was impressively increased up to a maximum of 88%, with significant clustering around the 80% range over the optimisation process, further validating the reproducibility of the yields achieved. Even though the optimal STY_{Diox} conditions were identified during the initial LHC, the MVMOO algorithm further went on to explore and map out the trade-off between the objectives to build five experimental Pareto points to highlight this.

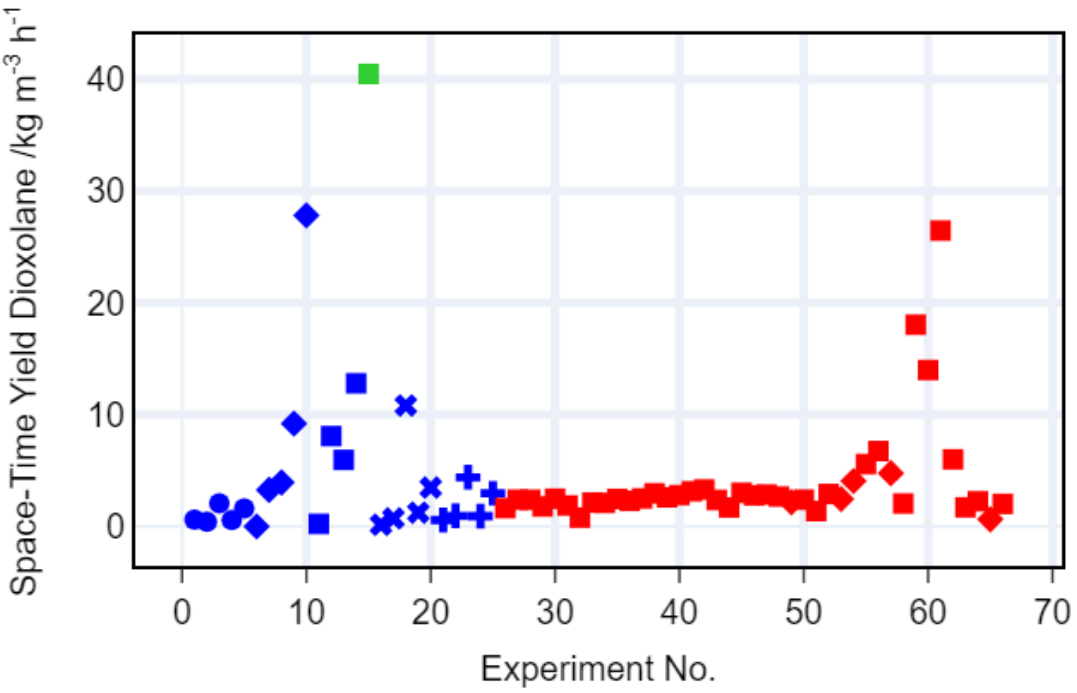


Figure 46. Results from the Heck-hydrolysis optimisation, with STY_{Diox} vs Experiment number plotted to highlight improvements made to this metric over the process. Where ■ – Initial, ■ – Algorithm and ■ – Optimum.

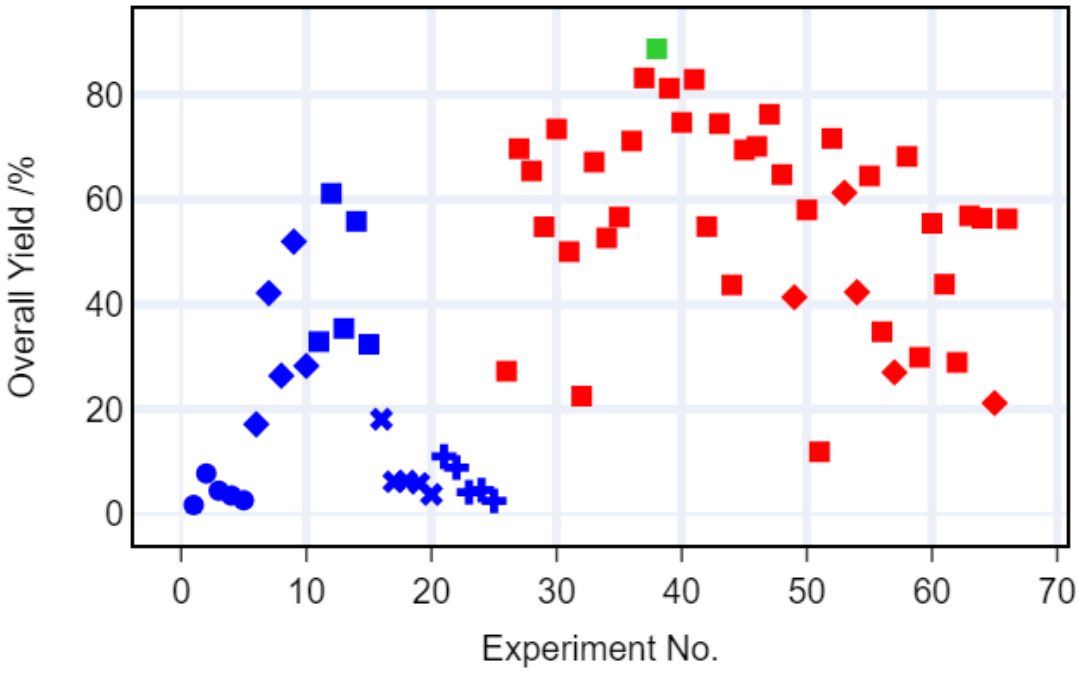


Figure 47. Results from the Heck-hydrolysis optimisation, with RME vs Experiment number plotted to highlight improvements made the RME metric over the process. Where ■ – Initial, ■ – Algorithm and ■ – Optimum.

Without the exploration and discovery of these experimental points, the information for the extent of the detrimental effects between the objectives would have been left unknown and the operating conditions and interactions between the variables to achieve these data points would have been unexploited. These results further justify the efficacy of the combined MVMOO-telescoped approach for exploring more productive regions in the objective space where alternative methods may have missed these optimal conditions and variable interactions.

It can be observed how the MVMOO algorithm was able to effectively map out the relative importance of each continuous variable when combined with dppp across the Pareto front, Figure 48. The maximisation of the overall yield in the telescoped process, required high residence times, maximum equivalents of the vinyl ether (**4.6**) and low temperatures, whereas the flow rate of the acid ratio variable was less important in this case. The temperature variable played a key role in this metric and understanding of the reaction pathway, as at lower temperatures the concentration of intermediate formation is increased over the dioxolane concentration.

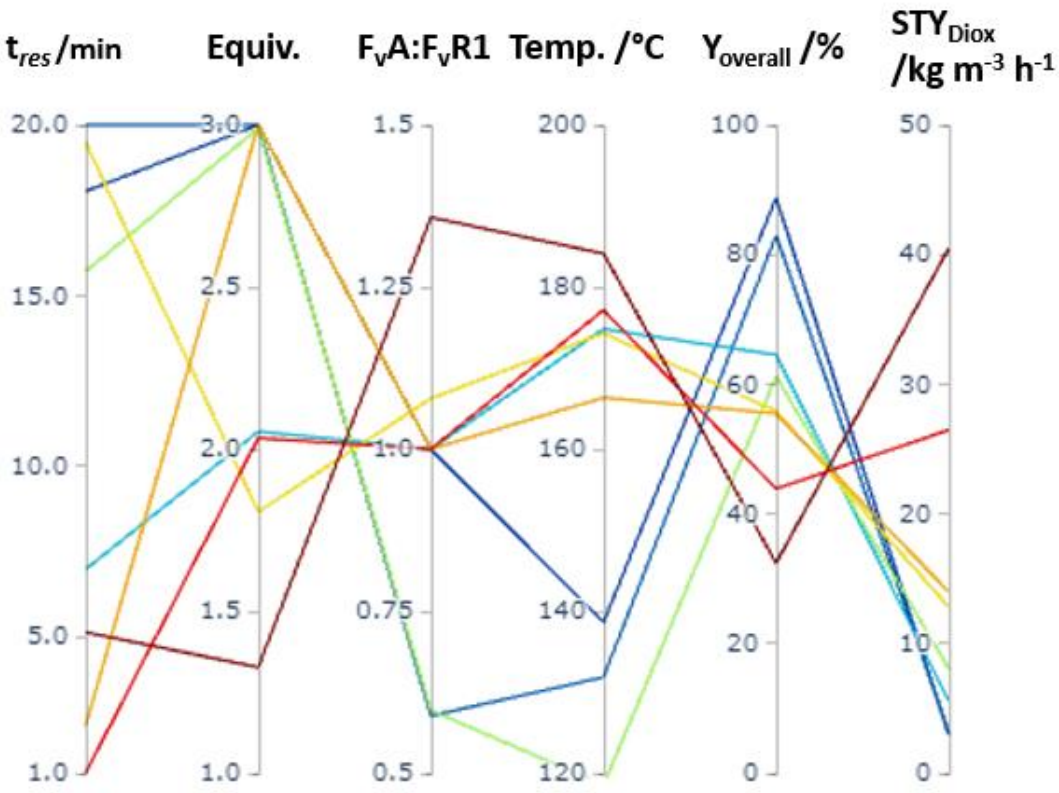


Figure 48. Parallel coordinate plot highlighting the interactions between continuous variables and objective outcomes for each Pareto optimal solution.

As previously mentioned, this intermediate is more readily hydrolysed to the final product, as such leading to higher overall yields observed, indicating the pivotal role of the temperature for this process metric. On the other hand, to achieve high STY_{Diox} can be achieved using low residence times, moderate-low equivalents and higher temperatures required, with the acid flow rate ratio not being examined for this objective due to it only participating in the hydrolysis reaction and having no effect on the dioxolane formation in the Heck reaction. The higher temperatures required further explain the temperature dependence on overall yields observed where elevated temperatures are essential to advance the cyclisation of the alpha intermediate for the formation of dioxolane.

Utilising the MVMOO algorithm in this telescoped optimisation enabled it to be completed in only 62 hours, requiring no human intervention, highlighting the efficiency and applicability of these types of workflows for telescoped reactions. Similarly, to the previously presented case studies, employing MVMOO for simultaneous optimisation of mixed variables enables

real-time monitoring of the hypervolume in the optimisation process to act as a measure of process efficiency over time, Figure 49.

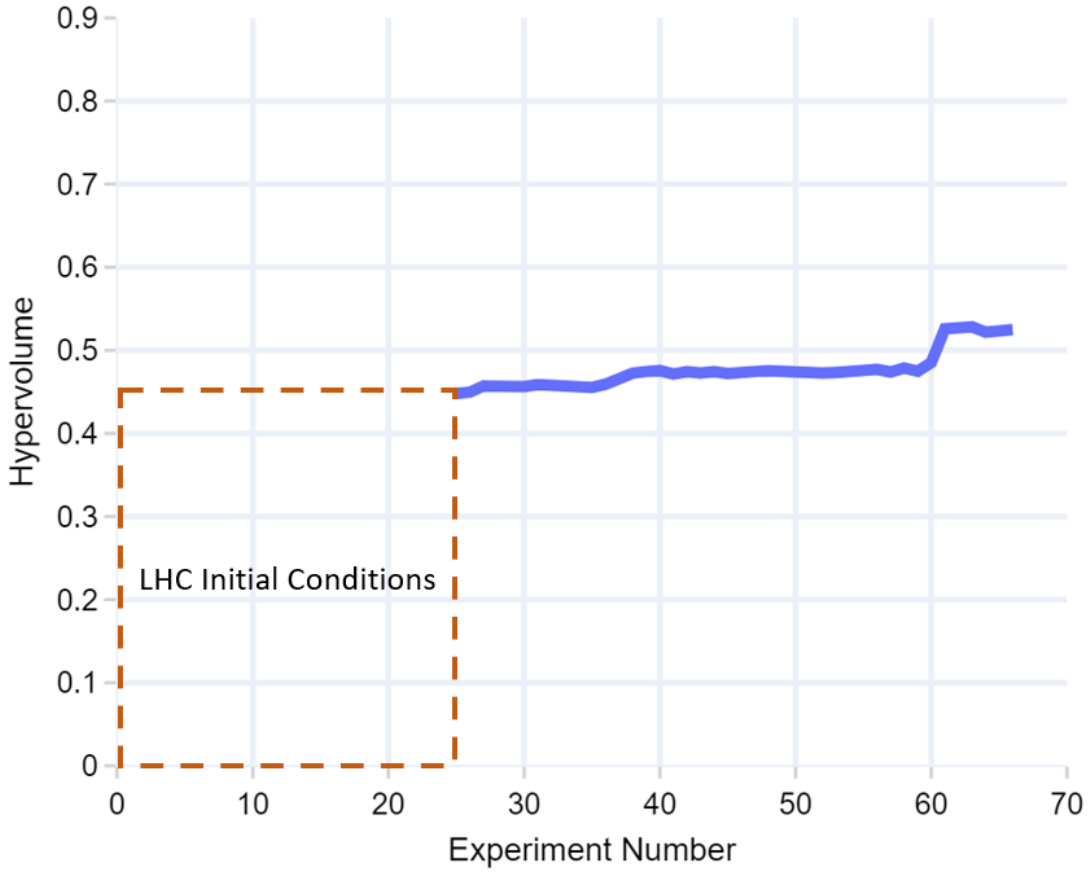


Figure 49. Hypervolume vs. Experiment number plot for the telescoped Heck-hydrolysis reaction that was monitored throughout the optimisation process.

This is critically important in the scope of telescoped optimisations where the associated cost per experiment is significantly increased, such that there is a necessity for the minimisation of experimental numbers. As before, once the experimental process had reached a minimum of 60 iterations, to ensure efficient exploration of the design space, the hypervolume would be monitored until a plateau for 5 experiments persisted. As it can be seen, due to the LHC identifying two exceptional points for the STY_{Diox} metric, the hypervolume forms an initial plateau whilst the algorithm is exploring high overall yielding experiments. After incremental hypervolume increases over the process, there is a significant increase around the 60th experiment related to two back-back Pareto optimal points exploring that trade-off between the metrics. After the 60th experiment, monitoring of the hypervolume revealed the presence of

a plateau, which then prompted the termination of the optimisation process to preserve material consumption for the process. This further justifies the methodology of real-time monitoring of the hypervolume, as termination before the increase observed at the 60th experiment would have led to a reduction in non-dominated solutions identified and information gained from the optimisation. Employing this approach facilitated the maximisation of information gained in the optimisation, whilst preserving material and time consumption for the process.

4.4 Conclusion

The heightened interest in telescoped optimisations of chemical systems over recent years presents the new focus for expanding automated reactor platforms onto complex multistage chemical synthesis.[20,145,152,154]

In this work, the methodology for optimising continuous variable telescoped systems presented by Clayton *et al*. was successfully advanced to introduce the inclusion of discrete variables within these processes.[20] This addition facilitated the increase in the number of variable interactions possible, subsequently leading to further process understanding between the interactions and reaction pathways occurring within the chemical system. Multi-point sampling after both reactors on the platform enabled the use of a single analytical HPLC technique for the quantification and analysis over both steps of the reaction under study.

To further explore the potential reaction pathway effects on the reaction outcome, the ligand selection was chosen as the discrete variable in this study, with both monodentate- and bidentate-phosphine ligands selected. The inclusion of a mixed set of phosphine ligands facilitated the examination of the pathway produced for the alkene insertion and the associated outcomes on the process metrics. These process metrics were set such that a multi-objective optimisation was achieved for each step in the reaction pathway. The objectives were assigned in a conflicting manner to explore the detrimental effects each objective had on one another. The first objective metric selected was the overall yield for the telescoped process. This selection was made to observe the operating conditions and critical interactions

required to produce a greater yielding system. For the Heck reaction, the objective was decided to be the STY for the dioxolane (**4.8**) formation. This choice stemmed from the known concept that the second hydrolysis reaction, to form the acetophenone (**4.9**), was much faster for the alpha-intermediate (**4.7**) than for the cyclised dioxolane, forcing the algorithm to explore the conditions required to map out the Pareto front between the objectives. The selection of both these objectives enabled the full exploration of the continuous variables throughout the process, where only low temperatures would have been studied if the STY of the alpha-intermediate was decided, providing a greater process within this optimisation. Additionally, this telescoped methodology included the set hypervolume stopping criteria that have been highlighted for each case study within this thesis. Inclusion of this stopping criteria based on the hypervolume progression over the optimisation is critical for these types of workflows, where the associated costs per experiment are significantly increased. This facilitated the optimisation to be completed in only 66 experiments, where maximisation of both process metrics was achieved.

In summary, the work in this chapter has highlighted the possibility of expanding the previous Chapter 2 and 3 workflows onto telescoped reactions. This work represents the completion of the first known multi-objective mixed variable optimisation on telescoped flow systems, providing enhanced efficiency and identification of key interactions over sequential reaction steps. This progress not only shows the potential for optimising mixed variable telescoped systems but provides further evidence for the applicability of these types of workflows for accelerating early-stage multi-step reaction development.

Chapter 5. Conclusions and Future Work

The work contained in this thesis has focused on the investigation and development of an automated flow platform with optimisation algorithms capable of handling mixed-variable systems for pharmaceutically relevant compounds. Later work in this thesis focused on the advancement of the capabilities of this platform to further highlight the applicability of these processes for early-stage reaction development.

The history of prior self-optimisation processes presented in the literature highlights the absence of mixed-variable multi-objective optimisations to comprehensively explore the full variable interactions within a system. Therefore, the inclusion of these parameters within these methodologies would further increase the efficiency and process understanding gained for chemical systems throughout these studies. Hence, the work for this thesis aimed at exploring: (i) the introduction of multi-objective optimisations for mixed-variable chemical systems; (ii) the development of these optimisations onto pharmaceutical relevant complex chemical pathways; (iii) the implementation of multi-step reactions for the consideration of mixed-variable multi-objective optimisations. This thesis has demonstrated the contribution to each of these respective areas through the optimisation of various relevant case studies.

The work in Chapter 2 described the simulated and experimental optimisations applying MVMOO, a Bayesian mixed-variable multi-objective algorithm previously developed by Jamie Manson, on an automated flow platform.[63] Initial work presents the potential insights that simulated optimisations provide as a precursor to experimental studies, utilising previously reported kinetic data coupled with SNOBFIT and a genetic algorithm to explore the objective space of a chemical system. Later studies investigated the development of the mixed-variable flow platform to facilitate the handling of discrete variables and explored the selection of the S_NAr reaction related to the known solvent effects on the regioselective outcome.[108,109,167] This approach selected two conflicting objectives for the optimisation based upon the solvent-dependant regioselective outcome, where the solvent of choice was designated the discrete variable to further explore the interaction effects required to evaluate the trade-off between

process metrics. The optimisation was performed concerning four continuous parameters in addition to the solvent choice, where the successful identification of the trade-off between the competing regioselective products during the process was achieved. This study underlined the inherent effects that the solvent polarity index had on the outcome of the reaction in addition to the continuous variable interactions required to maximise each objective. The prior simulated optimisation that was performed provided a rationale for the observations within the experimental study. Although this investigation highlighted the effectiveness of the efficient optimisation of mixed variable chemical systems, future work should focus on developing the system such that it is capable of a higher number of discrete variables. This can be achieved using a switching valve with a greater number of ports available, to ensure that the solvent effects of a wider array can be explored to further observe the underlying solvent properties that affect the reaction outcome. Additionally, the utilisation of multiple sample loops should be a focus for future work to enable multiple different discrete variables, bases, and solvents, within the optimisation pathway to explore the interactions and regioselective outcomes for these reaction pathways. Implementation of these additional discrete variables facilitates a greater process understanding gained from these optimisation procedures.

In Chapter 3 the potential for the mixed-variable multi-objective optimisations was further extended to include more complex catalytic systems, with the selection of pharmaceutically relevant Sonogashira cross-coupling being made. This reaction was selected based on the desired product being a precursor to a TRPV1 antagonist used within the pharmaceutical industry for pain treatment and management.[136,137] The inclusion of a catalytic system enabled an increase in the complexity of the optimisation aiming to explore the comprehensive variable interactions within these systems. To align with the evolving pathways of the pharmaceutical industry towards more environmentally sustainable API synthesis, two conflicting objectives were selected to explore the trade-off between productivity and environmental effects for reaction optimisation.[118] The optimisation was completed with respect to three continuous variables, with the selection of the ligand as the discrete variable due to its significant impact

on the outcomes of catalytic pathways. This study successfully identified the trade-off between RME and STY, notably highlighting the least sterically hindered ligand TPP, as the best-performing ligand, which contrasts with current chemical understanding.

Although hypervolume monitoring throughout the optimisation enabled the termination of the process once a plateau was observed, the consumption of material is a key consideration for optimisations, especially for processes involving catalytic systems where the associated costs of catalysts and resources are increased. Subsequently, future work should investigate the application of nanomole-scale high throughput experimentation combined with the presented self-optimisation platform to ensure the minimal consumption of material per iteration in the optimisation. The development of an HTE-flow platform would streamline the study of a higher number of discrete variables whilst eliminating the necessity for preparing individual stock solutions for each component. This, in turn, reduces the labour requirements and material consumption, both of which are key criteria for early-stage reaction development.

Multi-stage chemical synthesis for end-to-end processes in continuous flow is an area which can significantly benefit from the consideration of multi-factor optimisations. Observation of the synergistic effects between variables over multiple reaction pathways will provide greater process understanding and enhance efficiency. In Chapter 4, the work addressed the self-optimisation of mixed variable telescoped chemical systems with respect to multiple objectives, which have not been previously reported. Initial developments to the platform implemented multi-point sampling to achieve analysis and quantification over both steps on a single piece of analytical equipment, that was based on a previously reported technique.[20] This approach examined two competing objectives for both reaction pathways to explore the trade-off required within a Heck intramolecular cyclisation-hydrolysis telescoped reaction. Optimisation was completed concerning four continuous variables and one discrete parameter with the inclusion of continuous variables over both steps to examine the interactions for the entire reaction pathway. The ligand as the discrete variable enabled the influence of the potential catalytic pathways within the Heck reaction to be examined,

where a selection of mono and bidentate phosphine ligands was included to further explore the underlying mechanistic effects.[157] The Heck-hydrolysis telescoped reaction was successfully optimised in only 62 hours, requiring no human intervention, whilst achieving efficient exploration of the trade-off curve between the objectives. Therefore, this work presents the applicability of mixed-variable multi-objective optimisations towards the end-to-end processing of telescoped reactions and will likely be further extended to explore early-stage reaction development of APIs. Consequently, the future should aim to explore a higher number of reaction steps to further highlight the application of these optimisations towards process development.

The development of a nanomole-scale HTE-flow platform will aid in alleviating these challenges enabling access to a greater consideration of all factors throughout the telescoped pathway. However, to facilitate this advancement, future work will be required to include the consideration of discrete variables over all steps within the reaction pathways to examine a comprehensive understanding of the mixed variable interactions throughout all stages. Furthermore, the inclusion of multiple discrete factors for each reaction pathway will be required to further develop the process understanding gained from these optimisations.

To conclude, the development of automated flow platforms combined with optimisation algorithms in feedback loops has enabled the efficient exploration of chemical systems for the consideration of various objectives. The work in this thesis has focused on the development of mixed variable multi-objective optimisations to provide a comprehensive understanding of all factor interactions within chemical processes to further enhance efficient exploration of the trade-off between competing objectives. The identification required from future work to develop this area for early-stage reaction development has been highlighted to include: (i) reducing material consumption and labour requirements by implementing an automated nanomole-scale high throughput flow platform for mixed-variable multi-objective optimisations; (ii) increasing the number of discrete variables under consideration to further explore the underlying properties and interactions required to provide efficient optimisation; (iii) development of the telescoped optimisation to include a greater number of steps for process development; (iv) inclusion of multiple

discrete variables throughout all stages within a telescoped reaction pathway to grant greater process understanding from these optimisations. To facilitate the application of these mixed variable methodologies to the wider research community, the development of readily available automated control and optimisation documentation is critical for enabling ease of use to non-specialised chemists. The main focal point for the development of future work heavily relies on the implementation of nanomole-scale HTE with automated flow platforms which have become increasingly commercially available. However, the associated costs for the equipment currently remain a limitation for accessibility, but with the increasing interest in self-optimisation, the production of reduced-cost alternatives is inevitable. Therefore, providing accessibility of these technologies to the wider research community for the complete process optimisation and development of chemical systems.

Chapter 6. Experimental

6.1 Discrete Variable Automated Flow Platform

Reagents were made up to their desired concentrations in stock solutions during the experimental setup, this was completed by utilising spreadsheets to calculate the exact concentrations required. The solutions were loaded into conical flasks and primed on the dual piston reciprocating JASCO PU2800 and PU4185 HPLC pumps. Multiple stock solutions were made up for each discrete variable under study, where the ligands and solvent selection were achieved using a JASCO CO4062 column oven module installed with a 7-port 6-position switching valve. The discrete variable stock solutions were loaded onto each position of the selection valve, whereby the switching position of the valve would change the discrete variable selected, this solution was then fed into one of the HPLC pumps and supplied into the system. The streams from the HPLC pumps were mixed using Swagelok SS-100-3 tee-pieces in the orientation required for each different case study. Upon combination of all streams into a single flowing stream, it would be passed through a tubular reactor block of a desired volume. These reactors were made from either PTFE tubing (1/16” OD, 1/32” ID) or Stainless-steel tubing (1/16” OD, 1/32” ID), which were fitted in a tabular format to a cylindrical aluminium block and heated with a Eurotherm 3200 temperature controller to allow for the rapid heating of reaction liquids. Additionally, a conventional desk fan was equipped and pointed towards the heating block to aid in enhancing the cooling of the aluminium reactor when required. After the reactor, an aliquot of the reaction solutions was sampled for analysis in the feedback loop, which was achieved using a VICI Valco EUDA-CI4W sample loop (4-port) fitted with either a 0.02 µL or 0.5 µL sample loop injection volume. The sampled solution was fed directly into an Agilent 1260 Infinity II series HPLC instrument fitted with an Agilent Poroshell 120 EC-C18 reverse phase column (5 cm length, 4.6 mm ID and 2.7 µL particle size) for quantitative analysis to be performed, running a developed HPLC method for adequate signal separation. The flow system was maintained under a constant desired fixed back pressure using a 250 psi Upchurch Scientific back pressure regulator. The automated system was controlled through a custom-written MATLAB program to allow for real-time control and monitoring of all instruments involved in the system. Additionally, the MVMOO algorithm was employed

through Python, with the capability to communicate with MATLAB to enable the transfer of data between the algorithm and the control program. The data would include the analytical calculations for the experiments and the next set of conditions generated by the algorithm to be set on the equipment. To achieve the analytical calculations, internal standards (IS) would be included in the reservoir solutions to allow for the analysis of IS and compound signal areas to be evaluated for the desired metric calculations. Throughout the optimisation process, external monitoring of the optimisation process was completed using Microsoft Teams screen sharing capabilities. The fully annotated photo of the flow platform utilised for all work completed is shown below in Figure 50.

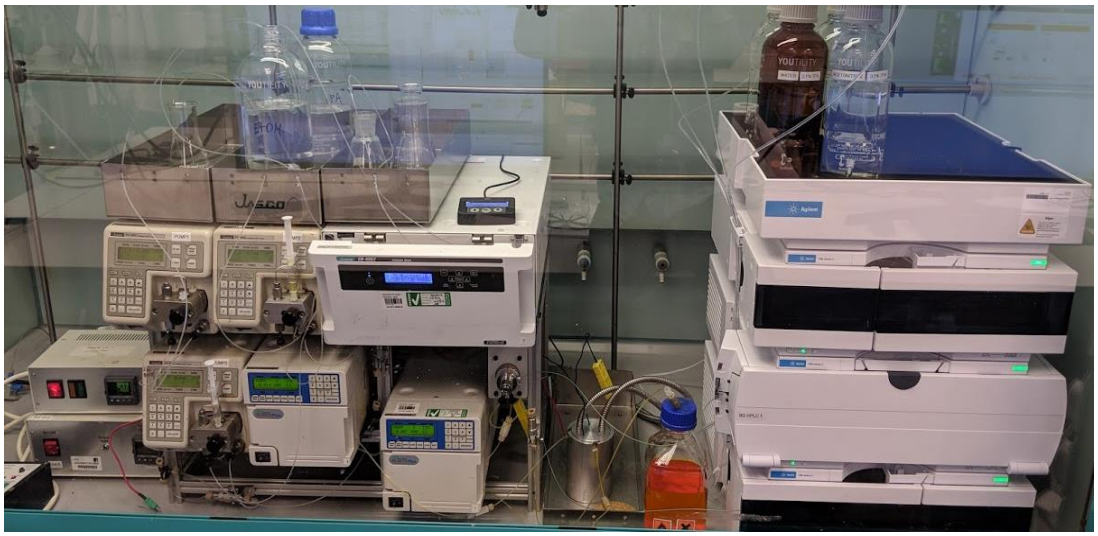


Figure 50. Photo of the automated mixed variable flow reactor used for Chapters 2, 3 and 4.

Controlling the pump flow rates, valve positions, reactor temperature, and sampling process was executed through a MATLAB script via RS232 control. During each iteration, valve positions were aligned with their corresponding discrete variables. The reactor was then permitted to stabilise at the target operational temperature. To conserve resources and expedite the process, pump flow rates were minimised during reactor heating/cooling, and initial LHC experiments were sequences based on ascending temperature values. Furthermore, sequential LHC experiments were initiated while the analysis of the preceding experiment was ongoing. Responses for each objective were ascertained from HPLC chromatograms obtained after each iteration. These responses played a pivotal role in updating the surrogate

models and generating subsequent reaction conditions using the MVMOO algorithm. Throughout the experimental sequence, the hypervolume was continually monitored. The optimisation process was halted when a substantial plateau in performance was observed, generally after around 60 experiments to ensure a sufficient number of points had been gathered.

6.2 Offline Analytical Equipment

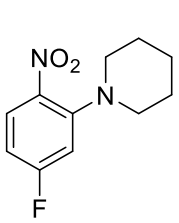
A Bruker 400 AVANCE III HD NMR Spectrometer (^{1}H NMR at 400 MHz, ^{13}C at 101 MHz) was used to perform NMR spectroscopy with the appropriate deuterated solvent. Chemical shifts in both ^{13}C and ^{1}H NMR spectra are reported as ppm downfield from TMS, and reported as singlet (s), doublet (d), triplet (t), quartet (q) and a combination for multiple instances, or multiplet (m). For instances with coupling present, coupling constants (*J*) are averaged between coupling signals and quoted in Hz. LC-MS analysis was achieved with an Agilent 1290 series uPLC and a Bruker HCT-Ultra detector with electrospray ionization (ESI) in the positive mode.

6.3 Chapter 2 Procedures

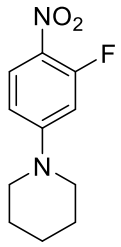
6.3.1 Chemicals

2,4-Difluoronitrobenzene **1** (99%, Fluorochem), morpholine **2** (99.0+%, Fisher Scientific Ltd.), triethylamine (99% Acros Organics), 1-methyl-2-pyrrolidone (NMP; 99%, Fisher Scientific Ltd.), dimethylformamide (DMF; Extra pure, Fisher Scientific Ltd.) ethanol (EtOH; 99.8%, Fisher Scientific Ltd.), acetonitrile (MeCN; HPLC grade, Fisher Scientific Ltd.) and biphenyl (99.5% GC, Merck Life Science UK Ltd.) were purchased from suppliers and used without further purification. Standards of 4-(5-fluoro-2-nitrophenyl)morpholine **3**, *4*-(3-fluoro-4-nitrophenyl)morpholine **4** and 4,4'-(4-nitro-1,3-phenylene)dimorpholine **5** were synthesised and characterised for HPLC calibration.

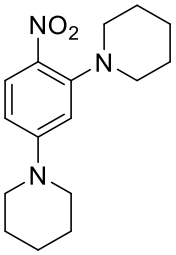
6.3.2 Synthesis of *ortho*-2.3, *para*-2.4 and bis-2.5



2.3

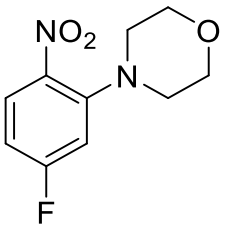


2.4

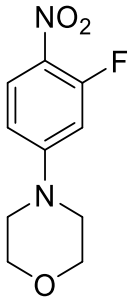


2.5

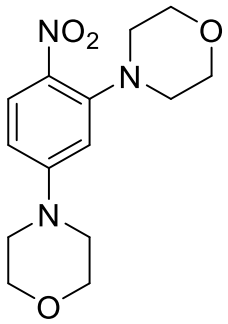
Morpholine **2.2** (6.02 g, 69.1 mmol) was added to 2,4-difluoronitrobenzene **2.1** (5.00 g, 31.4 mmol) in ethanol (150 mL) in a round-bottomed flask. The reaction mixture was left to stir at room temperature for 5 hours. The resultant mixture was concentrated *in vacuo*, redissolved in ethyl acetate (100 mL), and washed successively with saturated NH_4Cl solution (100 mL) and brine (100 mL). The organic layer was separated, dried over Na_2SO_4, and concentrated *in vacuo*. The resultant residue was purified by flash column chromatography (10-80% EtOAc/*n*-hexane) to afford *ortho*-**2.3** (5.53 g, 78%) as an orange oil, *para*-***2.4*** (0.8415 g, 12%) as a bright yellow solid and bis-adduct **2.5** as an orange solid (0.35 g, 4%).



^{1}H NMR ($CDCl_3$, 500 MHz) δ 7.90 (dd, *J* = 9.1, 6.0 Hz, 1H), 6.78 (dd, *J* = 11, 2.6 Hz, 1H), 6.72 (ddd, *J* = 9.5, 7.0, 2.6 Hz, 1H), 3.93 – 3.75 (m, 4H), 3.13 – 2.99 (m, 4H); ^{13}C NMR ($CDCl_3$, 126 MHz) δ 165.5 (d, *J* = 260 Hz), 148.6, 148.5, 129.0 (d, *J* = 11 Hz), 108.6 (d, *J* = 24 Hz), 107.2 (d, *J* = 25 Hz), 66.6, 51.7; *m*/*z* (ESI$^+$) $C_{10}H_{11}FN_2O_3$ $[M+H]^+$, calculated 227.08, found 227.33; in agreement with published data. [168]



^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.09 – 8.00 (t, *J* = 9.1 Hz, 1H), 6.61 (dd, *J* = 9.4, 2.7 Hz, 1H), 6.54 (dd, *J* = 15, 2.7 Hz, 1H), 3.89 – 3.82 (m, 4H), 3.40 – 3.33 (m, 4H); ^{13}C NMR ($CDCl_3$, 126 MHz) δ 158.0 (d, *J* = 260 Hz), 155.9, 155.8, 128.2 (d, *J* = 1.3 Hz), 108.2, 101.0 (d, *J* = 26 Hz), 66.2, 46.9; *m*/*z* (ESI$^+$) $C_{10}H_{11}FN_2O_3$ $[M+H]^+$, calculated 227.08, found 227.32; in agreement with published data. [168]



^{1}H NMR ($CDCl_3$, 500 MHz) δ 8.03 (d, *J* = 9.3 Hz, 1H), 6.47 (dd, *J* = 9.4, 2.6 Hz, 1H), 6.33 (d, *J* = 2.6 Hz, 1H), 3.91 – 3.87 (m, 4H), 3.87 – 3.82 (m, 4H), 3.37 – 3.30 (m, 4H), 3.10 – 3.03 (m, 4H); ^{13}C NMR ($CDCl_3$, 126 MHz) δ 155.2, 149.4, 133.1, 129.7, 107.0, 103.4, 66.9, 66.4, 52.3, 47.3; *m*/*z* (ESI$^+$) $C_{14}H_{19}N_3O_4$ $[M+H]^+$, calculated 294.15, found 294.39; in agreement with published data. [168]

6.3.3 Simulated Optimisation

Construction of concertation profiles for the simulated code was made based on previously reported reaction kinetic data, using numerical analysis of ODEs. Arrhenius kinetics facilitated the computation calculation of the corresponding rate constants for each transformation within the S_NAr pathway. Employing an optimisation algorithm onto these concentration profiles and kinetic models supplied by the ODEs enabled the refinement of the desired process metrics for the system.

$$\frac{dC_A}{dt} = -k_1 \cdot C_A \cdot C_B - k_2 \cdot C_A \cdot C_B \quad \textbf{(20)}$$

$$\frac{dC_B}{dt} = -k_1 \cdot C_A \cdot C_B - k_2 \cdot C_A \cdot C_B - k_3 \cdot C_B \cdot C_C - k_4 \cdot C_B \cdot C_D \quad \textbf{(21)}$$

$$\frac{dC_C}{dt} = k_1 \cdot C_A \cdot C_B - k_3 \cdot C_B \cdot C_C \quad \textbf{(22)}$$

$$\frac{dC_D}{dt} = k_2 \cdot C_A \cdot C_B - k_4 \cdot C_B \cdot C_D \quad \textbf{(23)}$$

$$\frac{dC_E}{dt} = k_3 \cdot C_B \cdot C_C + k_4 \cdot C_B \cdot C_D \quad \textbf{(24)}$$

Where C_A= Concentration of 2,4-difluoronitrobenzene (**2.1**), (M)

C_B = Concentration of morpholine (**2.2**), (M)

C_C = Concentration of *ortho* (**2.3**), (M)

C_D = Concentration of *para* (**2.4**), (M)

C_E = Concentration of *bis* (**2.5**), (M)

t = time, (min)

k_1 = Rate of reaction for the first step, (min^{-1})

k_2 = Rate of reaction for the second step, (min^{-1})

k_3 = Rate of reaction for the third step, (min^{-1})

k_4 = Rate of reaction for the fourth step, (min^{-1})

```
function [dydt] = SnArReact2( ~ , C0 , T )
%{
                Kinetics and catalytic data taken from:
        https://pubs.rsc.org/en/content/articlepdf/2017/re/c6re00109b
%}
% redefine concentrations
concA = C0(1); % initial conc of Nitrobenzene (A)
concB = C0(2); % initial conc of Morpholine (B)
concC = C0(3); % initial conc of Ortho (C)
concD = C0(4); % initial conc of Para (D)
concE = C0(5); % initial conc of Bis (E)


% define kinetic constants

TRef = 363.15; % reference temperature for the k values (Kelvin)
A1 = 1.422; % pre-exponential factor for step 1 (M-1 min-1)
A2 = 0.414; % pre-exponential factor for step 2 (M-1 min-1)
A3 = 0; % pre-exponential factor for step 3 (M-1 min-1)
A4 = 0.090; % pre-exponential factor for step 4 (M-1 min-1)

Ea1 = 38.2e3; % activation energy for step 1 (J/mol)
Ea2 = 32.9e3; % activation energy for step 2 (J/mol)
Ea3 = 0; % activation energy for step 3 (J/mol)
Ea4 = 40.8e3; % activation energy for step 4 (J/mol)

R = 8.3145; % gas constant (J K-1 mol-1)

% reparameterised Arrhenius equation:
k1 = A1*exp((-Ea1/R)*((1/T)-(1/TRef)));
k2 = A2*exp((-Ea2/R)*((1/T)-(1/TRef)));
k3 = A3*exp((-Ea3/R)*((1/T)-(1/TRef)));
k4 = A4*exp((-Ea4/R)*((1/T)-(1/TRef)));

dAdt = - (k1 .* concA .* concB) - (k2 .* concA .* concB);
dBdt = - (k1 .* concA .* concB) - (k2 .* concA .* concB) - (k3 .* concC .* concB) - (k4 .* concD .* concB);
dCdt = (k1 .* concA .* concB) - (k3 .* concC .* concB);
dDdt = (k2 .* concA .* concB) - (k4 .* concD .* concB);
dEdt = (k3 .* concC .* concB) + (k4 .* concD .* concB);


dydt = [dAdt dBdt dCdt dDdt dEdt]';
end

```

Figure 51. Screenshot of the concentration profiles used within the simulated code for the S_NAr reaction.

The simulated self-optimizations were conducted concerning four continuous parameters: t_{res}, pyrrolidine/morpholine **2.2** equivalents, concentration of **2.1** and temperature. The parameter limits are shown in Table 12. The single-objective optimization was to maximize *ortho* product **2.3** yield or maximise *para* product **2.4** yield, as defined by the [Eq (**25**-(**26**)].

minimize [-(*ortho* yield)] (**25**)

minimize [-(*para* yield)] (**26**)

```
function [Yield] = SnArFunction(X)
%To find yield of ortho product in the SnAr reaction
%   Inputting variables into the X value, this is later used in SNOBFIT for
%   the parameters

tRes = X(:,1);
equiv = X(:,2);
T = X(:,3) + 273.15;



concA = X(:,4);


concB = concA .* equiv; % Calculates the concentration of Morpholine
C0 = [concA concB zeros(1,3)]; % Indexes all 6 concentrations in C0

options = odeset('NonNegative', 1:length(C0));
options = odeset(options, 'RelTol', 1e-3);

[TimeData, ConcData] = ode45(@SnArReact, [0 tRes], C0, options, T);

CFinal = ConcData(end,:);

Output = - (100 * CFinal(4))/concA;

Yield = - Output;
%Productivity = CFinal(3)/tR;
%E factor =

end
```

Figure 52. Screenshot of the ODE Solver code required for the simulation of concentration profiles, returning the TimeData and Concentration Data, used to calculate the % yield results for each regioisomer in the simulation.

Table 6. Parameter boundaries for the four-variable single-objective simulated self-optimisation of the S_NAr reaction using pyrrolidine and morpholine in separate studies for comparison.

Limits	t_{res} /min	Pyrrolidine/Morpholine 2.2/equiv.	Conc. 2.1 /M	Temp /°C
Lower	0.5	1.0	0.05	60
Upper	2.0	5.0	0.175	120

Table 7. List of operating conditions and results from the simulated SNOBFIT optimisation of the S_NAr reaction for pyrrolidine with *ortho* yield as the objective. The optimal yield and conditions are highlighted in green.

Entry	t_{res}/min	Pyrrolidine Equiv.	Conc. 2.1/M	Temp/°C	*Ortho*-2.3 yield /%
1	1.3	2.3	0.3	99	68.6
2	0.5	3.4	0.2	60	90.7
3	1.4	4.5	0.3	120	14.6
4	0.5	3.8	0.3	80	82.7
5	0.6	2.1	0.38	110	76.3
6	0.7	1.1	0.4	70	82.9
7	0.7	4.9	0.29	90	65.2
8	1.5	4.9	0.35	105	20.7
9	1.4	1.2	0.15	85	87.2
10	1.1	4.9	0.43	66	66.9
11	1.8	4.6	0.42	74	47.2
12	0.9	1.1	0.22	115	88.5
13	0.4	4.5	0.1	60	82.6
14	0.5	1	0.12	93	66.4
15	0.4	1	0.27	90	74.8
16	1	1	0.1	73	64.2
17	0.7	2.4	0.21	95	85.1
18	1.8	5	0.1	60	86.1
19	1.9	1	0.25	93	87.1
20	0.4	2.2	0.33	60	87.1
21	1.6	1.7	0.41	85	80.4
22	1.8	1	0.38	60	84.2

23	1.9	2.2	0.11	91	84.5
24	0.6	2.5	0.23	80	89.2
25	1.8	2.7	0.46	97	45.1
26	1.8	2.1	0.3	60	87.6
27	1.9	1.6	0.22	80	87.9
28	0.4	1	0.32	101	80.5
29	0.9	3.9	0.22	115	51.8
30	1.8	2.6	0.26	60	85.7
31	1.9	1	0.12	81	80.5
32	2	2.2	0.21	85	78.6
33	1.3	4.2	0.29	65	76.5
34	0.4	3.5	0.32	60	90.7
35	0.4	1.7	0.25	85	88.0
36	0.4	2.4	0.41	84	87.6
37	1.4	2.2	0.16	112	73.9
38	0.7	3.1	0.29	60	89.9
39	0.9	1	0.1	104	77.6
40	0.4	1.7	0.5	81	90.1
41	0.4	1	0.24	82	69.0
42	1.7	2.3	0.29	82	76.5
43	1.3	2.3	0.3	99	68.6

Table 8. List of operating conditions and results from the simulated SNOBFIT optimisation of the S_NAr reaction for pyrrolidine with *para* yield as the objective. The optimal yield and conditions are highlighted in green.

Entry	t_{res}/min	Pyrrolidine Equiv.	Conc. 2.1/M	Temp/°C	*para*-2.4 yield /%
1	0.9	4.9	0.3	87	2.16
2	1.5	1.0	0.3	120	4.13
3	1.0	1.2	0.3	60	3.68
4	1.3	2.9	0.2	103	2.23
5	2	3	0.12	74	3.49
6	1.2	1	0.26	95	3.93
7	1.6	4	0.36	112	0.18
8	0.6	4.2	0.16	67	3.87
9	1.3	1.9	0.11	81	3.99
10	1.3	4.6	0.1	98	2.52
11	1	1.7	0.48	108	2.69
12	0.7	4.4	0.23	117	1.21
13	1.5	1	0.1	74	3.24
14	0.4	2.6	0.1	60	2.86
15	1	1	0.1	66	2.63
16	1.1	1	0.21	120	4.09
17	0.8	3.4	0.35	93	2.28
18	2	1	0.5	87	3.99
19	1.3	5	0.16	60	3.46
20	1.9	1	0.21	119	4.11
21	0.8	1.5	0.38	88	3.89
22	2	1	0.1	110	4.00

23	1	2.5	0.1	89	3.91
24	0.4	1	0.39	80	3.43
25	0.7	2.6	0.42	71	3.58
26	1.4	1	0.1	120	4.02
27	0.5	1.3	0.1	82	3.00
28	0.4	5	0.28	74	3.54
29	0.9	1	0.42	95	3.97
30	1.4	2.7	0.43	107	0.97
31	1.9	1	0.1	120	4.07
32	0.5	1	0.29	114	3.97
33	1.8	1.5	0.14	85	3.97
34	1.5	1.3	0.1	94	4.02
35	0.8	3.4	0.15	112	2.54
36	0.5	1	0.5	120	4.09
37	1.8	2.8	0.1	85	3.49
38	1.1	1	0.1	90	3.45
39	1.4	4.4	0.38	93	0.77
40	2	1	0.17	120	4.11
41	1.9	1.9	0.12	98	3.52
42	0.4	1	0.4	85	3.56
43	0.4	5	0.1	74	3.96
44	1.2	1.7	0.44	98	2.97

Table 9. List of operating conditions and results from the simulated SNOBFIT optimisation of the S_NAr reaction for morpholine with *ortho* yield as the objective. The optimal yield and conditions are highlighted in green.

Entry	t_{res}/min	Morpholine Equiv. 2.2	Conc. 2.1/M	Temp/°C	*Ortho*-2.3 yield /%
1	1.9	2.1	0.1	92	42.58
2	0.6	4.9	0.3	60	27.56
3	0.7	2.7	0.3	120	69.15
4	1.7	4.5	0.5	76	73.50
5	2	3.8	0.19	106	76.47
6	1.7	1.2	0.19	68	18.80
7	1.8	1.4	0.4	84	50.19
8	0.9	4.3	0.37	113	78.28
9	0.5	4	0.45	99	66.53
10	1.3	5	0.46	88	76.24
11	0.8	1.1	0.45	73	21.91
12	1.5	4.8	0.36	65	61.52
13	0.5	5	0.12	120	57.53
14	2	5	0.33	120	79.70
15	1.5	5	0.5	120	79.70
16	1.3	5	0.5	115	79.36
17	1.6	4.4	0.24	83	67.46
18	1.7	4.1	0.4	113	79.20
19	2	5	0.35	94	77.60
20	2	5	0.5	88	77.21
21	1	1.8	0.39	105	60.95
22	1.5	5	0.31	120	79.69

23	1.8	4.8	0.1	113	74.50
24	1.9	5	0.48	79	75.83
25	1.5	2.2	0.35	102	69.93
26	1.5	4.5	0.4	120	79.70
27	2	3.4	0.48	107	78.75
28	2	4.8	0.35	113	79.22
29	2	4.8	0.45	110	79.01
30	0.9	3.9	0.38	75	54.90
31	2	2.7	0.5	120	79.68
32	2	4	0.5	111	79.08
33	2	4.5	0.45	120	79.70
34	1	2	0.28	81	36.61
35	2	2.7	0.38	120	79.57
36	2	4.8	0.35	103	78.45
37	2	2.7	0.45	97	76.54
38	1.1	3.9	0.23	101	69.53
39	2	3.6	0.42	120	79.70
40	2	3.9	0.5	102	78.39
41	2	3.8	0.5	118	79.57
42	1	1.8	0.19	105	44.32

Table 10. List of operating conditions and results from the simulated SNOBFIT optimisation of the S_NAr reaction for morpholine with *para* yield as the objective. The optimal yield and conditions are highlighted in green.

Entry	**t_{res}/min**	**Morpholine Equiv.**	**Conc. 2.1/M**	**Temp/°C**	***para*-2.4 yield /%**

1	1.7	4.3	0.4	69	19.92
2	0.7	1.2	0.3	120	10.31
3	0.9	2.5	0.3	94	14.37
4	1.5	5.0	0.2	82	17.79
5	1.6	1.4	0.42	107	16.50
6	1	2.1	0.14	60	4.01
7	1.9	5	0.44	114	10.87
8	1.3	1.1	0.48	75	10.24
9	1.4	4.3	0.22	101	18.44
10	1.5	3.8	0.44	88	19.43
11	0.5	4.5	0.28	64	8.67
12	1.1	1	0.35	112	11.91
13	2	2.4	0.14	112	16.58
14	2	5	0.46	90	16.86
15	2	1	0.3	91	11.71
16	2	5	0.33	97	16.82
17	0.9	2.3	0.35	68	10.16
18	2	3.5	0.5	76	20.48
19	1.2	4.9	0.5	83	19.81
20	2	5	0.4	66	21.23
21	1	3.9	0.16	111	16.70
22	2	5	0.14	73	16.80
23	2	4.8	0.31	78	20.30
24	2	5	0.36	77	20.35
25	1.7	5	0.1	95	16.98
26	1	4.2	0.22	71	13.08

27	2	4.7	0.5	60	21.62
28	1.1	5	0.24	99	18.55
29	2	4.7	0.5	72	20.72
30	1	1.8	0.39	101	15.39
31	0.7	3.8	0.28	71	11.38
32	2	3.5	0.32	65	18.23
33	1.5	3.4	0.5	103	17.13
34	1.4	1.7	0.28	76	10.94
35	0.6	3.2	0.24	64	6.70
36	2	4.7	0.5	79	19.45
37	2	4.9	0.38	72	20.84
38	2	5	0.5	60	21.79
39	1.5	2.8	0.18	84	14.01
40	1.5	3.3	0.36	67	16.96
41	2	5	0.5	66	21.43
42	1.6	5	0.5	76	20.35
43	2	4.3	0.41	68	20.87
44	1.3	4.2	0.42	101	17.66

6.3.4 Experimental Set-Up

Reservoir solutions were prepared by dissolving the desired reagents in a solvent under stirring at ambient conditions until a homogenous reaction mixture was achieved. These solutions were then loaded and primed into the desired pumps for the optimisation.

Table 11. List of reservoir solutions for the S_NAr optimisation. Biphenyl was used as the internal standard for the optimisation process.

Reservoir	Compound 1	Compound 2	Solvent
1	2,4-Difluoronitrobenzene (27.79 mL, 0.25 mol, 1.015 mol L^{-1}) **R1**	Biphenyl (1.95 g, 13.00 mmol, 0.051 mol L^{-1}) **I.S**	The solvent of choice (NMP, DMF, DMAc, MeCN, EtOH) (250 mL)
2	Morpholine (45.58 mL, 1.04 mol, 2.085 mol L^{-1}) **R2**	**X**	Triethylamine (250 mL)
Solvent	**X**	**X**	The solvent of choice (NMP, DMF, DMAc, MeCN, EtOH)

The automated reactor was set up according to the schematic shown in Figure 54, where the reactor volume = 3 mL and the fixed back pressure = 100 psi. HPLC mobile phases were A H_2O (18.2 MΩ), and B MeCN, both buffered with 0.1% TFA. The method used was 10% to 90% B 5.0 mins, 90% to 10% B 0.1 mins, 10% B 1 min, flow rate 1.50 mL min^{-1}, column temperature 20 °C. An example chromatogram is shown in Figure 53.

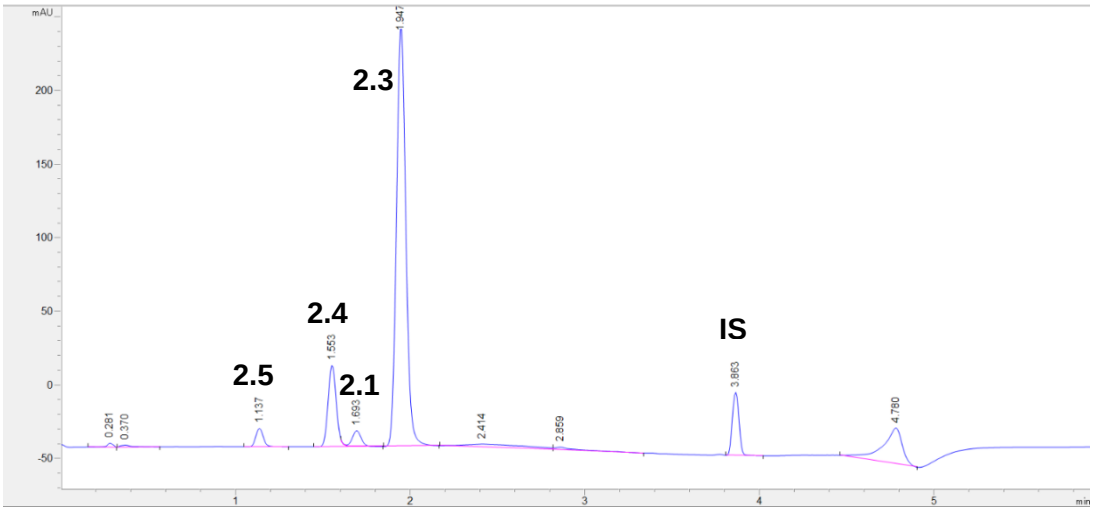


Figure 53. Example HPLC chromatogram for the S_NAr optimisation. Retention times (min): bis 5 = 1.14; para 4 = 1.53; 2,4-difluoronitrobenzene 1 = 1.69; ortho 3 = 1.95; biphenyl (IS) = 3.86

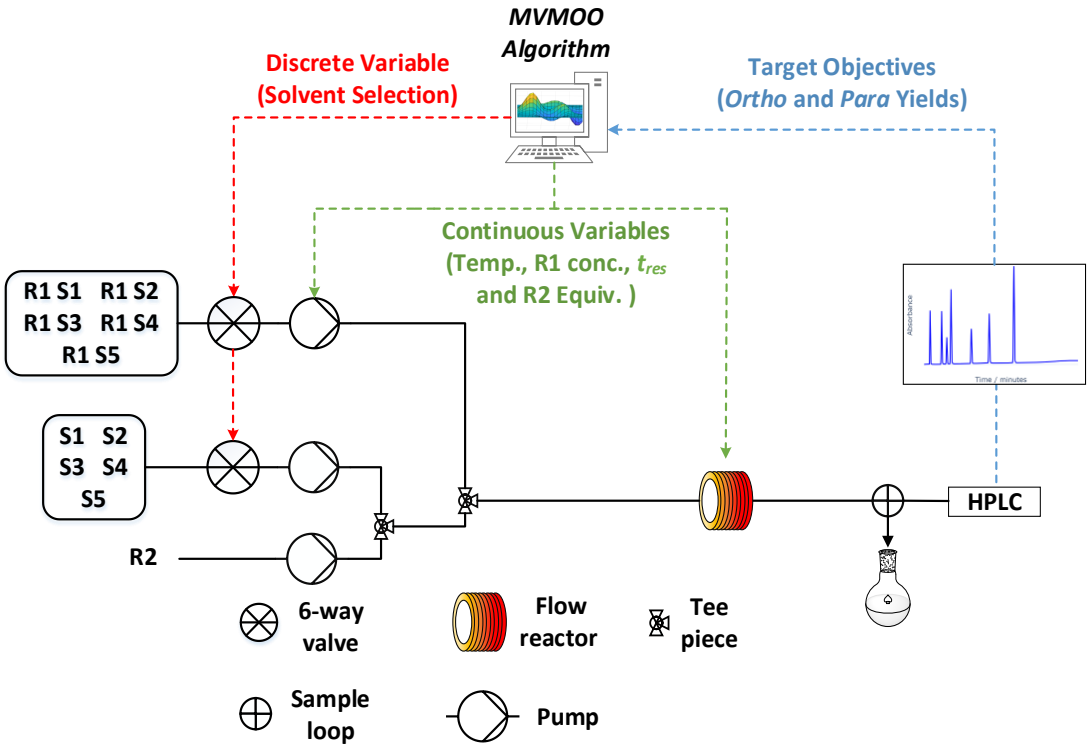


Figure 54. Flow schematic of the optimisation platform used for this chapter's work. R1 represents 2,4-difluoronitrobenzene **2.1** and R2 is morpholine **2.2.** Biphenyl (internal standard) was included in all the R1 stock solutions. S1 is DMF, S2 is NMP, S3 is EtOH, S4 is MeCN and S5 is DMAc.

6.3.5 Self-Optimisation Results

The self-optimization was conducted concerning four continuous parameters: t_{res}, morpholine **2.2** equivalents, concentration of **2.1** and temperature and one discrete parameter: valve position (solvent choice). The parameter limits are shown in Table 12. The objective of the optimization was to simultaneously maximize *ortho* product **2.3** yield and maximise *para* product **2.4** yield, as defined by the [Eq (**27**)].

minimise [-(*ortho* yield), -(*para* yield)] (**27**)

Table 12. Parameter boundaries for the five-variable experimental multi-objective self-optimisation of the S_NAr reaction using the MVMOO algorithm.

Limits	**t_{res} /min**	**Morpholine 2.2/equiv.**	**Conc. 2.1 /M**	**Temp /°C**	**Valve Position**
Lower	0.5	1.0	0.05	60	1
Upper	2.0	5.0	0.175	120	5

Table 13. List of operating conditions and results from the self-optimisation of the S_NAr reaction. The first 25 experiments were completed as a LHC design. Pareto optimal points are highlighted in green.

Entry	**t_{res}/min**	**Equiv. 2.2**	**Conc. 2.1/M**	**Temp/°C**	**Solvent**	***Ortho*-2.3 yield /%**	***Para*-2.4 yield /%**
1	1.19	2.3	0.146	62.7	DMF	44.3	28.9
2	1.19	2.3	0.146	62.7	NMP	39.6	35.5
3	1.19	2.3	0.146	62.7	EtOH	13.6	3.0
4	1.19	2.3	0.146	62.7	DMAc	34.4	28.8

5	1.19	2.3	0.146	62.7	MeCN	34.2	6.5
6	1.65	5.0	0.062	79.3	DMF	57.1	33.3
7	1.65	5.0	0.062	79.3	NMP	50.0	41.4
8	1.65	5.0	0.062	79.3	EtOH	31.0	6.7
9	1.65	5.0	0.062	79.3	DMAc	47.5	37.5
10	1.65	5.0	0.062	79.3	MeCN	56.3	8.5
11	0.56	4.0	0.098	90.8	DMF	45.1	26.7
12	0.56	4.0	0.098	90.8	NMP	40.0	33.3
13	0.56	4.0	0.098	90.8	EtOH	19.3	3.0
14	0.56	4.0	0.098	90.8	DMAc	35.6	29.0
15	0.56	4.0	0.098	90.8	MeCN	33.9	4.2
16	1.97	2.6	0.123	107.1	DMF	62.2	32.3
17	1.97	2.6	0.123	107.1	NMP	54.1	40.5
18	1.97	2.6	0.123	107.1	EtOH	59.6	7.9
19	1.97	2.6	0.123	107.1	DMAc	55.6	38.1
20	1.97	2.6	0.123	107.1	MeCN	74.1	10.2
21	0.83	1.6	0.157	112.1	DMF	49.2	28.8
22	0.83	1.6	0.157	112.1	NMP	44.1	35.6
23	0.83	1.6	0.157	112.1	EtOH	31.4	4.0
24	0.83	1.6	0.157	112.1	DMAc	41.8	32.4
25	0.83	1.6	0.157	112.1	MeCN	45.4	5.6
26	2.00	2.7	0.149	107.1	NMP	52.8	41.6
27	2.00	2.8	0.175	107.6	NMP	52.9	41.5
28	1.67	3.7	0.175	109.1	NMP	53.9	39.5
29	2.00	5.0	0.050	120.0	MeCN	79.8	10.4
30	2.00	3.7	0.050	117.1	DMAc	53.6	40.3

31	2.00	5.0	0.050	120.0	NMP	53.1	40.8
32	2.00	5.0	0.050	88.8	NMP	49.5	48.2
33	1.81	4.5	0.160	94.9	NMP	54.4	40.3
34	1.77	3.7	0.141	88.8	NMP	52.2	43.3
35	1.78	4.5	0.053	88.1	NMP	48.2	42.7
36	1.98	4.3	0.175	88.0	NMP	53.4	41.9
37	2.00	5.0	0.050	120.0	DMF	62.0	31.4
38	2.00	5.0	0.175	91.3	DMF	59.8	33.1
39	2.00	5.0	0.050	96.8	DMAc	51.7	42.3
40	2.00	1.0	0.050	85.9	DMF	28.6	18.5
41	1.95	5.0	0.175	108.2	DMF	59.8	29.2
42	2.00	1.0	0.050	120.0	NMP	19.0	15.9
43	2.00	3.3	0.175	98.5	NMP	52.9	42.2
44	2.00	5.0	0.097	113.5	DMAc	54.5	39.2
45	2.00	5.0	0.050	103.9	NMP	49.3	42.2
46	2.00	3.6	0.050	106.1	DMAc	45.3	37.7
47	1.85	3.6	0.175	120.0	DMF	59.5	29.7
48	1.52	5.0	0.175	120.0	DMAc	56.3	34.2
49	0.96	5.0	0.175	120.0	NMP	58.5	35.5
50	2.00	5.0	0.175	60.8	NMP	53.6	40.9
51	2.00	2.8	0.175	117.8	DMAc	55.6	38.0
52	1.39	5.0	0.175	119.0	NMP	56.2	35.7
53	2.00	4.0	0.171	94.1	NMP	53.7	41.6
54	2.00	5.0	0.175	66.2	DMF	59.6	36.6
55	1.46	2.7	0.175	104.2	NMP	52.7	42.1
56	2.00	5.0	0.175	71.8	NMP	54.5	41.8

57	0.72	5.0	0.050	120.0	DMF	53.3	30.3
58	2.00	3.4	0.175	105.4	DMF	59.4	32.6
59	0.80	5.0	0.050	116.3	NMP	50.2	41.2
60	0.50	5.0	0.050	120.0	EtOH	24.7	0.0
61	1.92	4.6	0.057	119.8	DMAc	54.2	41.3
62	1.56	5.0	0.175	119.2	DMF	61.4	28.0
63	2.00	4.4	0.132	120.0	DMAc	55.1	35.1
64	2.00	1.6	0.058	119.3	MeCN	52.6	7.2
65	1.03	4.8	0.080	120.0	NMP	52.9	43.1
66	2.00	5.0	0.175	113.3	MeCN	75.2	16.9
67	2.00	5.0	0.174	120.0	MeCN	79.0	14.0
68	2.00	2.5	0.164	112.8	DMAc	54.4	40.2
69	2.00	5.0	0.110	91.1	NMP	53.0	43.3
70	2.00	5.0	0.175	76.6	DMF	61.1	34.8
71	1.76	2.7	0.175	111.1	DMAc	54.5	40.1
72	2.00	2.3	0.175	105.7	DMAc	53.3	41.5
73	1.98	3.0	0.050	112.6	DMF	58.2	33.8
74	2.00	2.8	0.167	86.7	NMP	51.6	43.5
75	1.83	2.4	0.171	98.0	NMP	51.9	43.0
76	2.00	3.9	0.149	74.8	NMP	52.0	43.4
77	2.00	5.0	0.078	68.6	NMP	50.7	44.3
78	2.00	5.0	0.175	90.1	DMAc	55.3	40.4
79	2.00	4.5	0.169	113.3	MeCN	76.8	16.8
80	2.00	3.9	0.175	115.0	MeCN	75.9	17.6
81	2.00	5.0	0.142	77.4	NMP	53.7	41.4
82	2.00	5.0	0.175	71.2	DMF	59.9	37.0

83	2.00	5.0	0.077	93.6	NMP	51.5	40.3
84	2.00	3.5	0.105	88.1	NMP	51.7	43.5
85	1.86	2.6	0.166	102.5	NMP	52.8	42.1
86	2.00	4.6	0.139	115.5	MeCN	78.1	16.3
87	1.05	5.0	0.173	110.9	NMP	56.0	38.4
88	1.40	5.0	0.175	109.6	NMP	55.8	38.0
89	2.00	5.0	0.175	99.1	DMAc	55.4	38.4
90	2.00	4.3	0.099	67.8	NMP	50.3	44.1
91	1.46	5.0	0.175	68.6	DMF	59.2	35.2
92	1.40	5.0	0.175	64.9	NMP	52.3	41.5
93	2.00	5.0	0.175	120.0	DMF	61.1	25.9
94	1.55	3.2	0.134	110.2	DMF	60.5	33.9
95	1.05	4.8	0.163	115.8	NMP	55.5	39.2
96	1.73	5.0	0.050	60.1	DMF	46.4	29.1
97	2.00	4.9	0.170	66.7	NMP	53.4	43.0
98	1.84	2.7	0.175	109.4	DMF	60.1	34.6
99	1.71	4.9	0.050	115.3	DMF	59.8	35.8

Table 14. List of Pareto front experiments in order of optimum *ortho*-**2.3** yield % to optimum *para*-**2.4** yield %.

Entry	**t_{res}/min**	**Equiv. 2.2**	**Conc. 2.1/M**	**Temp/°C**	**Solvent**	***Ortho*-2.3 yield /%**	***Para*-2.4 yield /%**
1	2.00	5.0	0.050	120.0	MeCN	79.8	10.4
2	2.00	5.0	0.174	120.0	MeCN	79.0	14.0
3	2.00	4.6	0.139	115.5	MeCN	78.1	16.3

4	2.00	4.5	0.169	113.3	MeCN	76.8	16.8
5	2.00	3.9	0.175	115.0	MeCN	75.9	17.6
6	1.97	2.6	0.123	107.1	DMF	62.2	32.3
7	2.00	5.0	0.175	76.6	DMF	61.1	34.8
8	2.00	5.0	0.175	71.2	DMF	59.9	37.0
9	1.05	5.0	0.173	110.9	NMP	56.0	38.4
10	1.05	4.8	0.163	115.8	NMP	55.5	39.2
11	2.00	5.0	0.175	90.1	DMAc	55.3	40.4
12	2.00	5.0	0.175	71.8	NMP	54.5	41.8
13	2.00	4.9	0.170	66.7	NMP	53.4	43.0
14	2.00	5.0	0.110	91.1	NMP	53.0	43.3
15	1.77	3.7	0.141	88.8	NMP	52.2	43.3
16	2.00	3.9	0.149	74.8	NMP	52.0	43.4
17	2.00	3.5	0.105	88.1	NMP	51.7	43.5
18	2.00	2.8	0.167	86.7	NMP	51.6	43.5
19	2.00	5.0	0.078	68.6	NMP	50.7	44.3
20	2.00	5.0	0.050	88.8	NMP	49.5	48.2

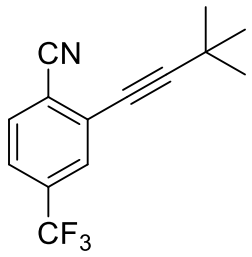
6.4 Chapter 3 Procedures

6.4.1 Chemicals

2-Bromo-4-(trifluoromethyl)benzonitrile **3.4** (95%, Fluorochem), 3,3-dimethyl-1-butyne **3.5** (98%, Merck Life Science UK Ltd.), palladium (II) acetate (99+%, Merck Life Science UK Ltd.), copper (I) iodide (98%, Merck Life Science UK Ltd.), 2-dicyclohexylphosphino-2'-(N,N-dimethylamino)biphenyl (DavePhos; 98%, Fluorochem), 2-dicyclohexylphosphino-2'.4'.6'-triisopropylbiphenyl (XPhos; 98%, Fluorochem), (2-biphenyl)dicyclohexylphosphine (CyJohnPhos; 98%,

Fluorochem), 2-dicyclohexylphosphino-2',6'-dimethoxybiphenyl (SPhos; 98%, Fluorochem), 2-methyl-2'-dicyclohexylphosphinobiphenyl (TPP; Flake 99%, Alfa Aesar), 1,3,5-trimethoxybenzene (99+%, Merck Life Science UK Ltd.), pyrrolidine (99%, Fisher Scientific Ltd.), acetonitrile (MeCN; HPLC grade, Fisher Scientific Ltd.) and toluene (PhMe; 99.8+%, Fischer Scientific Ltd.) were purchased from suppliers and used without further purification. A standard of the desired product **3.3** was synthesised and characterised.

6.4.2 Synthesis of 2-(3,3-dimethylbut-1-yn-1-yl)-4-(trifluoromethyl)benzonitrile 3.6



3.6

Pd_2dba_3 (1.11 mg, 1.22×10^{-3} mmol), DavePhos (1.90 mg, 4.85×10^{-3} mmol) and CuI (0.46 mg, 2.40×10^{-3} mmol) were added to a round-bottomed flask. The flask was then purged with nitrogen and degassed trimethylamine (2 mL) was added. 2-chloro-4-(trifluoromethyl)benzonitrile (0.07 mL, 0.49 mmol) was added and the reaction mixture was heated to 65 °C. 3,3-dimethyl-1-buytne **7** (0.01 mL, 0.79 mmol) was added slowly over 2 hours using a syringe pump. The reaction mixture was heated for an additional 2 hours. The resultant mixture was diluted with isopropylacetate (20 mL) and then washed twice with water (2×30 mL) and twice with 10% citric acid (2×30 mL). The collected organic layer was diluted with methanol (40 mL) and concentrated to ~5 mL *in vacuo*. This was repeated twice more using methanol (2×70 mL), then concentration fully *in vacuo*. The resultant residue was purified by flash column chromatography (0-20% EtOAc/*n*-hexane) to afford the desired product **3.6** as a pale-yellow oil (90.3 mg, 74%).

^{1}H NMR ($CDCl_3$, 400 MHz) δ 7.73 (d, *J* = 4.8 Hz, 2H), 7.58 (d, *J* = 8.4 Hz, 1H), 1.37 (s, 9H); ^{13}C NMR ($CDCl_3$, 101 MHz) δ 132.76, 128.67, 123.87, 118.63, 116.32, 107.98, 74.90, 30.33, 28.22; *m/z* (ESI$^+$) $C_{14}H_{12}F_3N$ $[M+H]^+$, calculated 251.09, found 251.38; in agreement with published data.[136]

6.4.3 Experimental Set-Up

Reservoir solutions 1 and 2 were prepared by dissolving the desired reagents in a solvent under stirring at ambient conditions. Reservoir solutions 3 (a-e) were prepared under inert conditions and sonicated at 40 °C in a water bath until a homogenous solution was achieved.

Table 15. List of reservoir solutions for the Sonogashira optimisation. 1,3,5-trimethoxybenzene was used as the internal standard for the optimisation process.

Reservoir	**Compound 1**	**Compound 2**	**Compound 3**	**Compound 4**	**Solvent**
1	2-Bromo-4-(trifluoromethyl)-benzonitrile (24.00 g, 0.096 mol, 0.6 mol L^{-1})	1,3,5-Trimethoxybenzene (4.04 g, 0.024 mol, 0.15 mol L^{-1}) **I.S**	**X**	**X**	Toluene/ MeCN (2:1, 160 mL)
2	3,3-Dimethyl-1-butyne (16.6 mL, 0.14 mol, 0.400 mol L^{-1})	**X**	**X**	**X**	Toluene/ MeCN (2:1; 240 mL)
3A	DavePhos (0.76 g, 1.9 mmol, 0.012 mol L^{-1})	Palladium acetate (0.14 g, 0.64 mmol, 0.004 mol L^{-1})	CuI (1.83 g, 9.6 mmol, 0.06 mol L^{-1})	Pyrrolidine (20.48 g, 0.288 mol, 1.8 mol L^{-1})	Toluene/ MeCN (2:1, 160 mL)

3B	XPhos (0.92 g, 1.9 mmol, 0.012 mol L^{-1})	Palladium acetate (0.14 g, 0.64 mmol, 0.004 mol L^{-1})	CuI (1.83 g, 9.6 mmol, 0.06 mol L^{-1})	Pyrrolidine (20.48 g, 0.288 mol, 1.8 mol L^{-1})	Toluene/ MeCN (2:1, 160 mL)
3C	CyJohnPhos (0.67 g, 1.9 mmol, 0.012 mol L^{-1})	Palladium acetate (0.14 g, 0.64 mmol, 0.004 mol L^{-1})	CuI (1.83 g, 9.6 mmol, 0.06 mol L^{-1})	Pyrrolidine (20.48 g, 0.288 mol, 1.8 mol L^{-1})	Toluene/ MeCN (2:1, 160 mL)
3D	SPhos (0.79 g, 1.9 mmol, 0.012 mol L^{-1})	Palladium acetate (0.14 g, 0.64 mmol, 0.004 mol L^{-1})	CuI (1.83 g, 9.6 mmol, 0.06 mol L^{-1})	Pyrrolidine (20.48 g, 0.288 mol, 1.8 mol L^{-1})	Toluene/ MeCN (2:1, 160 mL)
3E	Triphenylphosphine (0.50 g, 1.9 mmol, 0.012 mol L^{-1})	Palladium acetate (0.14 g, 0.64 mmol, 0.004 mol L^{-1})	CuI (1.83 g, 9.6 mmol, 0.06 mol L^{-1})	Pyrrolidine (20.48 g, 0.288 mol, 1.8 mol L^{-1})	Toluene/ MeCN (2:1, 160 mL)
Solvent	**X**	**X**	**X**	**X**	Toluene/MeCN (2:1)

The automated reactor was set up according to the schematic shown in Figure 55, where the reactor volume = 4 mL and the fixed back pressure = 250 psi. HPLC mobile phases were A H_2O (18.2 MΩ), and B MeCN, both buffered with 0.1% TFA. The method used was 40% B 1.5 mins, 40 to 95% B 2.5 mins, 95% to 40% B 0.1 min, 40% B 1 min, flow rate 1.50 $mLmin^{-1}$, column temperature 20 °C. An example chromatogram is shown in Figure 56.

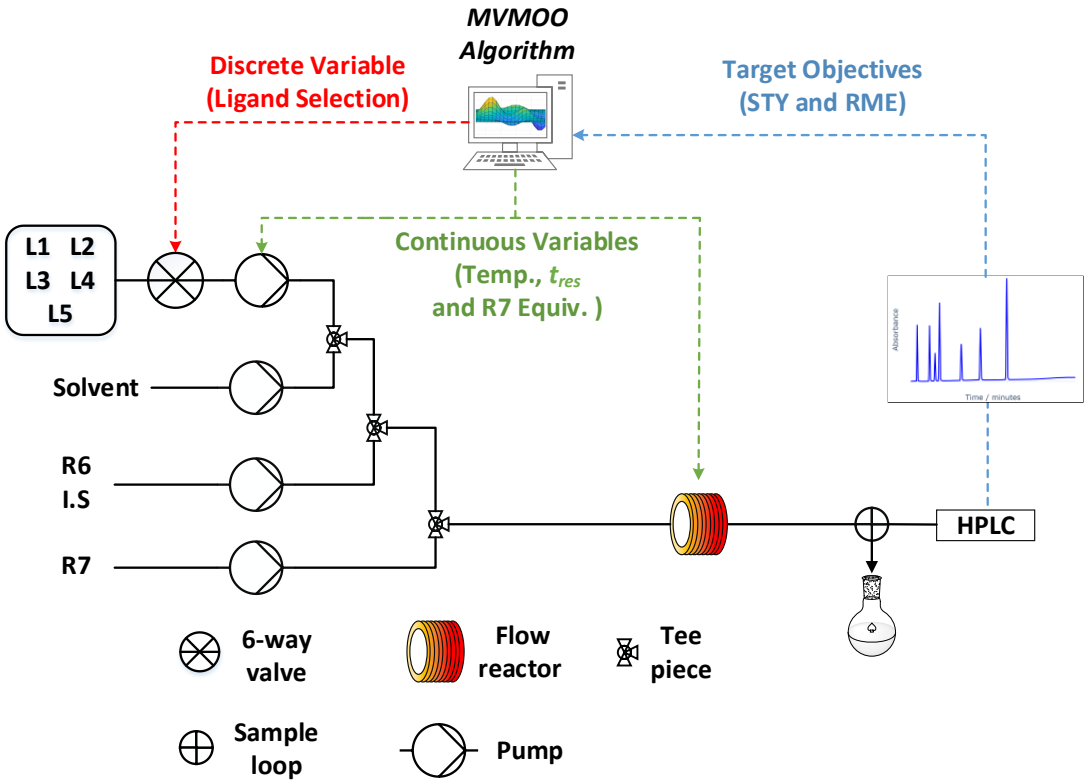


Figure 55. Flow schematic used for the Sonogashira optimisation. R6 represents 2-bromo-4-(trifluoromethyl) benzonitrile **3.4**, I.S is the internal standard 1,3,5-trimethoxybenzene and R7 represents 3,3-dimethylbutyne **3.5**. L1 is DavePhos, L2 is XPhos, L3 is CyJohnPhos, L4 is SPhos and L5 is TPP. The catalyst $Pd(OAc)_2$, CuI and pyrrolidine base was included in each of these stock solutions.

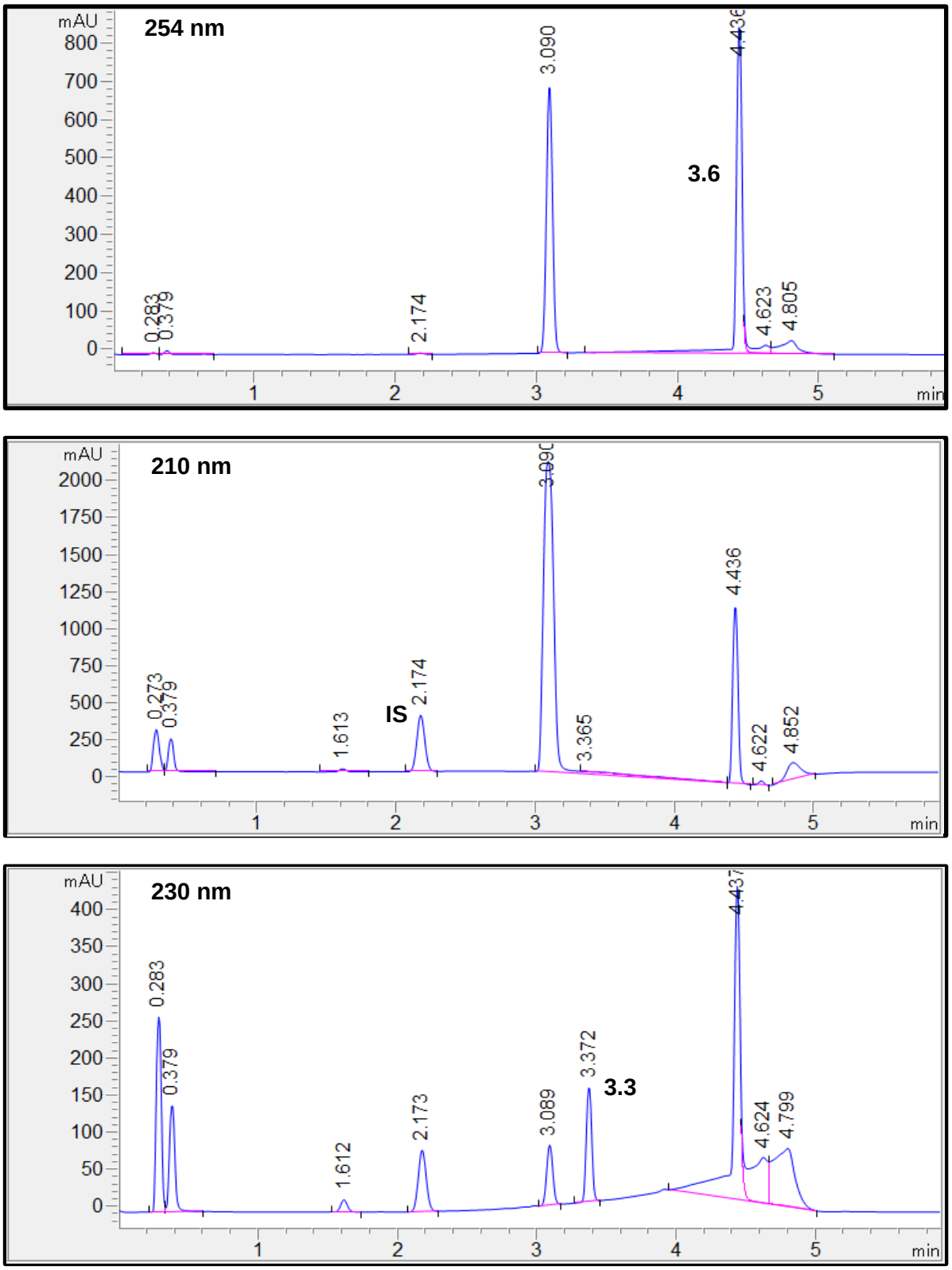


Figure 56. Example HPLC chromatogram for the Sonogashira reaction. Retention times (min): 2-chloro-4-(trifluoromethyl)benzonitrile **3.4** = 3.37 (230 nm); 2-(3,3-dimethylbut-1-yn-1-yl)-4-(trifluoromethyl)benzonitrile **3.6** = 4.44 (254 nm); 1,3,5-trimethoxybenzene (IS) = 2.17 (210 nm).

6.4.4 Self-Optimisation Results

The self-optimization was conducted concerning three continuous parameters: t_{res}, 3,3-dimethylbutyne **3.5** equivalents and temperature and one discrete parameter: valve position (ligand choice). The parameter limits are shown in Table S2. The objective of the optimization was to simultaneously maximize space-time yield (STY) as defined by [Eq (**28**)] and maximise the reaction mass efficiency (RME) [Eq (**29**)] of the reaction as defined by the [Eq (**30**)]

$$STY = \frac{mass_{product}}{Volume \times t_{res}} \quad (\mathbf{28})$$

$$RME = \frac{MW_{product\ \mathbf{3.6}} \times Yield}{MW_{\mathbf{3.4}} + (MW_{\mathbf{3.5}} \times equiv_{\mathbf{3.5}})} \quad (\mathbf{29})$$

$$\text{minimise [-(RME), -(STY)]} \quad (\mathbf{30})$$

Table 16. Parameter boundaries for the four-variable self-optimisation of the Sonogashira coupling reaction.

Limits	**t_{res}**	**Eq. 3.5**	**Temp /°C**	**Valve position**
Lower	1.0	1.0	60	1
Upper	10.0	3.0	140	5

Table 17. List of operating conditions and results from the self-optimisation of the Sonogashira coupling reaction. The first 25 experiments were completed as a LHC design. Pareto optimal points are highlighted in green.

Entry	**t_{res}/min**	**Equiv. 3.2**	**Temp/°C**	**Ligand**	**RME**	**STY/kg m^{-3} h^{-1}**
1	9.3	1.86	63.5	DavePhos	18.77	9.06
2	3.6	2.53	76.9	DavePhos	37.06	54.91

3	1.9	1.07	99.7	DavePhos	32.18	60.69
4	7.6	2.99	119.5	DavePhos	3.22	2.57
5	6.2	1.66	137.3	DavePhos	40.33	27.18
6	9.3	1.86	63.5	XPhos	21.30	10.40
7	3.6	2.53	76.9	XPhos	15.28	22.87
8	1.9	1.07	99.7	XPhos	30.59	58.50
9	7.6	2.99	119.5	XPhos	39.82	32.11
10	6.2	1.66	137.3	XPhos	40.54	27.65
11	9.3	1.86	63.5	CyJohnPhos	14.52	6.97
12	3.6	2.53	76.9	CyJohnPhos	15.22	22.42
13	1.9	1.07	99.7	CyJohnPhos	26.01	48.71
14	7.6	2.99	119.5	CyJohnPhos	45.39	36.08
15	6.2	1.66	137.3	CyJohnPhos	56.89	38.10
16	9.3	1.86	63.5	SPhos	54.82	26.53
17	3.6	2.53	76.9	SPhos	45.46	67.48
18	1.9	1.07	99.7	SPhos	38.45	72.73
19	7.6	2.99	119.5	SPhos	45.86	36.70
20	6.2	1.66	137.3	SPhos	56.87	38.42
21	9.3	1.86	63.5	TPP	68.19	32.31
22	3.6	2.53	76.9	TPP	59.08	86.08
23	1.9	1.07	99.7	TPP	52.00	95.94
24	7.6	2.99	119.5	TPP	50.42	39.66
25	6.2	1.66	137.3	TPP	62.97	41.62
26	2.1	1.80	62.7	TPP	48.71	116.87
27	2.4	2.77	73.8	TPP	56.61	140.43
28	3.6	2.43	139.9	TPP	51.56	78.98

29	6.8	2.92	72.3	TPP	55.09	49.45
30	9.2	1.35	71.2	TPP	63.98	31.41
31	6.4	1.00	136.8	TPP	52.43	34.20
32	1.9	3.00	82.3	TPP	47.66	151.41
33	1.9	3.00	66.2	TPP	44.84	147.63
34	2.0	3.00	92.0	TPP	52.52	161.10
35	1.0	3.00	140.0	SPhos	51.51	322.05
36	2.4	2.81	140.0	CyJohnPhos	46.76	115.89
37	7.9	1.00	61.3	TPP	53.15	27.98
38	2.3	2.81	140.0	SPhos	46.85	124.97
39	5.1	2.07	85.2	TPP	58.87	60.07
40	1.6	2.00	87.2	TPP	52.98	167.98
41	1.6	1.67	126.6	TPP	57.71	178.37
42	1.4	1.72	128.0	TPP	57.27	193.29
43	1.3	1.72	128.0	TPP	55.90	203.91
44	1.2	1.62	129.2	TPP	53.61	211.10
45	1.2	1.92	120.6	TPP	53.89	237.84
46	10.0	1.76	134.8	TPP	54.48	26.80
47	1.0	1.87	133.4	TPP	53.94	271.17
48	6.0	2.22	74.0	TPP	63.17	56.87
49	1.0	3.00	140.0	TPP	45.70	280.86
50	1.3	1.00	139.7	SPhos	41.47	135.19
51	7.4	2.11	72.4	TPP	64.63	46.07
52	4.9	2.42	75.4	TPP	59.26	66.90
53	4.0	1.71	124.0	TPP	57.30	69.60
54	1.2	1.56	138.6	TPP	52.59	205.00

55	9.2	2.36	72.3	TPP	61.44	37.03
56	1.0	2.65	103.0	TPP	48.68	282.19
57	1.2	1.68	127.3	TPP	52.93	206.21
58	4.8	2.23	74.0	TPP	63.19	70.82
59	8.3	2.02	65.3	TPP	65.55	41.01
60	1.3	1.75	127.3	TPP	54.12	208.28
61	4.2	2.30	73.5	TPP	62.10	80.12
62	3.2	2.31	73.8	TPP	62.19	104.72
63	1.3	2.37	74.6	TPP	41.42	175.59
64	2.4	2.10	108.4	TPP	53.56	115.76
65	6.7	2.05	60.2	TPP	63.25	48.95
66	3.0	2.42	74.6	TPP	59.44	111.82
67	10.0	1.00	140.0	SPhos	27.38	11.72
68	4.5	2.06	62.4	TPP	63.42	73.72
69	1.4	1.96	134.0	TPP	54.86	198.67

Table 18. List of Pareto front experiments in order of optimum RME to optimum STY.

Entry	**t_{res}/min**	**Equiv. 7**	**Temp/°C**	**Ligand**	**RME**	**STY/kg m^{-3} h^{-1}**
1	9.29	1.86	63.5	TPP	68.19	32.31
2	8.28	2.02	65.3	TPP	65.55	41.01
3	7.39	2.11	72.4	TPP	64.63	46.07
4	4.49	2.06	62.4	TPP	63.42	73.72
5	3.25	2.31	73.8	TPP	62.19	104.72
6	2.96	2.42	74.6	TPP	59.44	111.82

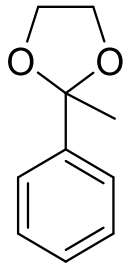
7	1.57	1.67	126.6	TPP	57.71	178.37
8	1.45	1.72	128.0	TPP	57.27	193.29
9	1.34	1.72	128.0	TPP	55.90	203.91
10	1.28	1.75	127.3	TPP	54.12	208.28
11	1.00	1.87	133.4	TPP	53.94	271.17
12	1.00	3.00	140.0	SPhos	51.51	322.05

6.5 Chapter 4 Procedures

6.5.1 Chemicals

Bromobenzene **4.5** (99%, Alfa Aesar), Ethylene glycol vinyl ether **4.6** (98%, Merck Life Science UK Ltd.), acetophenone **4.9** (98%, Fluorochem), palladium (II) acetate (99+%, Merck Life Science UK Ltd.), triethylamine (99% Acros Organics), bis(diphenylphosphino)methane (dppm: 97%, Fluorochem), 1,2-bis(diphenylphosphino)ethane (dppe: 95%, Fluorochem), 1,3-Bis(diphenylphosphino)propane (dppp: 95%, Fluorochem), 2-dicyclohexylphosphino-2'-(N,N-dimethylamino)biphenyl (DavePhos; 98%, Fluorochem), 2-methyl-2'-dicyclohexylphosphinobiphenyl (TPP; Flake 99%, Alfa Aesar), nitric acid (69-72%, Fisher Scientific Ltd.), acetonitrile (MeCN; HPLC grade, Fisher Scientific Ltd.) and ethane diol (99+%, Fisher Scientific Ltd.) were purchased from suppliers and used without further purification. A standard of the desired dioxolane product **4.8** was synthesised and characterised.

6.5.2 Synthesis of 2-methyl-2-phenyl-1,3-dioxolane 4.8



4.8

An oven-dried, two-necked round-bottom flask containing a stirrer bar was charged with bromobenzene (0.16 g, 1.0 mmol), $Pd(OAc)_2$ (2.25 mg, 0.01 mmol), dppp (8.25 mg, 0.2 mmol) and EG:MeCN (2:1) solvent (2 mL) under nitrogen at room temperature. Following degassing three times, an ethylene glycol vinyl ether (0.26 g, 0.28 mL, 3.0 mmol) and NEt_3 (0.25 g, 0.34 mL, 2.5 mmol) were sequentially injected. The flask was placed in an oil bath, and the mixture was stirred and heated at 145 °C for 2 hours. After this reaction period, the flask was removed from the oil bath and cooled to room temperature. The resultant mixture was diluted with H_2O (30 mL) the aqueous phase was extracted with Et_2O (3 x 20 mL) and the combined organic layers were washed with H_2O (20 mL). The combined organic phase was dried over Na_2SO_4 and then concentrated *in vacuo*. The resultant mixture was passed through a silica gel-filled Pasteur pipette using DCM as the eluent. The product was concentrated *in vacuo*, to yield the dioxolane (**4.8**) product as a pale yellow crystalline solid (0.1445 g, 88%). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.41-7.39 (m, 2 H), 7.26-7.18 (m, 3 H), 3.97-3.94 (m, 2H), 3.71-3.66 (m, 2 H), 1.58 ppm (s, 3 H); ^{13}C NMR (101 MHz, $CDCl_3$): δ 143.7, 128.6, 128.3, 125.7, 109.3, 64.9, 28.1 ppm; *m/z* (ESI$^+$) $C_{10}H_{12}O_2$ $[M+H]^+$, calculated 165.08, found 165.10; in agreement with published literature.[165]

6.5.3 Experimental Set-Up

Reservoir solutions 1 (a-e) were prepared under inert conditions and sonicated at 50°C until a homogenous solution was achieved. Reservoir solutions 2 and 3 were prepared by dissolving the desired reagents in the solvent under stirring at ambient conditions.

The automated reactor was set up according to the schematic shown in Figure 55, where the reactor volume = 3 mL and the fixed back pressure = 250 psi. HPLC mobile phases were A H2O (18.2 MΩ) and B MeCN. The method used was 16.3% B 2 min, 16.3 to 95.0% B 7 min, 95.0% B 1 min, 95.0 to 16.3% B 0.1 min, 16.3% B 0.9 min, flow rate 2.0 mL min-1 , column temperature 40 °C. In this case, the same method was used for both reaction steps, resulting in a total analysis time of 15 min for the telescoped process. Example chromatograms are shown in Figure 57.

Table 19. List of reservoir solutions for the Heck-hydrolysis telescoped optimisation. Methyl p-tolyl sulfone was used as the internal standard for the optimisation process.

Reservoir	Compound 1	Compound 2	Compound 3	Compound 4	Compound 5	Solvent
1A	Bromobenzene (7.85 g, 0.05 mol)	$Pd(OAc)_2$ (0.11 g, 0.5 mmol)	dppm (0.38 g, 1.0 mmol)	NEt_3 (12.65 g, 17.4 mL, 0.125 mol)	Methyl p-tolyl sulfone (3.41 g, 0.02 mol) **IS**	EG/ MeCN (1:1, 100 mL)
1B	Bromobenzene (7.85 g, 0.05 mol)	$Pd(OAc)_2$ (0.11 g, 0.5 mmol)	Dppe (0.38 g, 1.0 mmol)	NEt_3 (12.65 g, 17.4 mL, 0.125 mol)	Methyl p-tolyl sulfone (3.41 g, 0.02 mol) **IS**	EG/ MeCN (1:1, 100 mL)
1C	Bromobenzene (7.85 g, 0.05 mol)	$Pd(OAc)_2$ (0.11 g, 0.5 mmol)	Dppp (0.41 g, 1.0 mmol)	NEt_3 (12.65 g, 17.4 mL, 0.125 mol)	Methyl p-tolyl sulfone (3.41 g, 0.02 mol) **IS**	EG/ MeCN (1:1, 100 mL)
1D	Bromobenzene (7.85 g, 0.05 mol)	$Pd(OAc)_2$ (0.11 g, 0.5 mmol)	TPP (0.26 g, 1.0 mmol)	NEt_3 (12.65 g, 17.4 mL, 0.125 mol)	Methyl p-tolyl sulfone (3.41 g, 0.02 mol) **IS**	EG/ MeCN (1:1, 100 mL)
1E	Bromobenzene (7.85 g, 0.05 mol)	$Pd(OAc)_2$ (0.11 g, 0.5 mmol)	DavePhos (0.39 g, 1.0 mmol)	NEt_3 (12.65 g, 17.4 mL, 0.125 mol)	Methyl p-tolyl sulfone (3.41 g, 0.02 mol) **IS**	EG/ MeCN (1:1, 100 mL)

2	Ethylene glycol vinyl ether (17.62, 18.9 mL g, 0.2 mol)	**X**	**X**	**X**	**X**	EG/ MeCN (1:1, 250 mL)
3	HNO_3 (20.19 g, 14.3 mL 0.90 M)	**X**	**X**	**X**	**X**	Water (250 mL)
Solvent	**X**	**X**	**X**	**X**	**X**	EG/ MeCN (1:1, 100 mL)

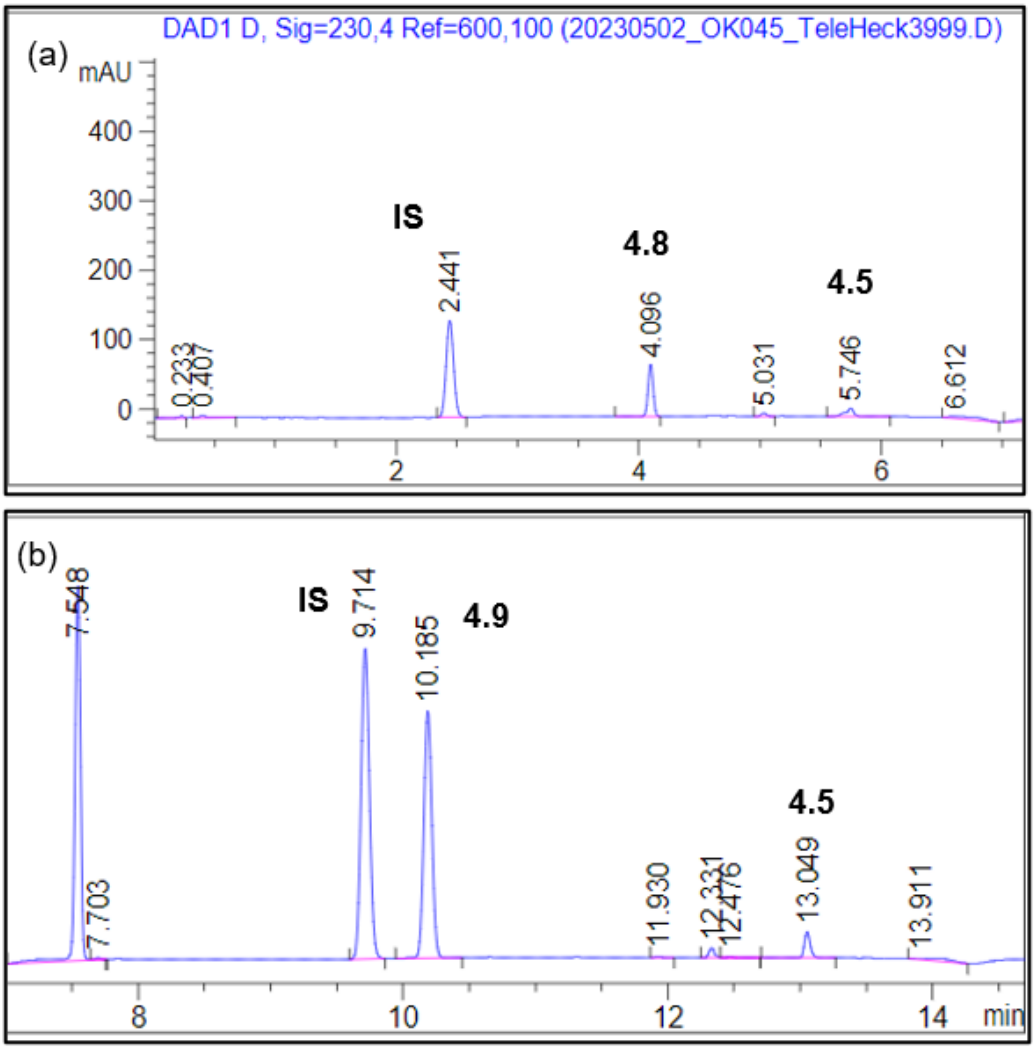


Figure 57. Example HPLC chromatogram at 230 nm for the Heck-hydrolysis reaction. (a) Heck-intramolecular cyclisation (first step); (b) Hydrolysis deprotection (second step). Retention times (min): Bromobenzene **4.5** = 5.74 and 13.05 min; Dioxolane **4.8** = 4.09 and 11.93 min; acetophenone **4.9** = 2.92 and 10.19 min and methyl p-tolyl sulfone (IS) = 2.44 and 9.71 min.

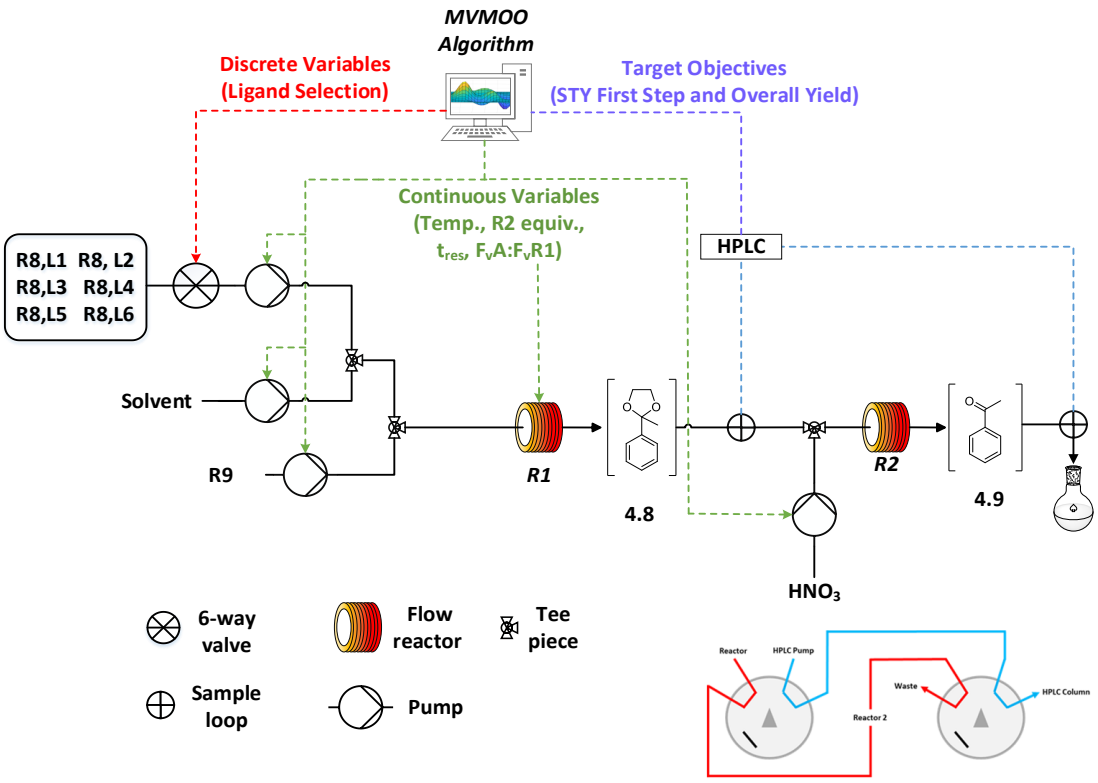


Figure 58. Flow schematic used for the Heck-hydrolysis telescoped optimisation. R8 represents bromobenzene **4.5** and R9 represents ethylene glycol vinyl ether **4.6**. L1 is dppm, L2 is dppe, L3 is dppp, L4 is DavePhos and L5 is TPP. The catalyst $Pd(OAc)_2$, triethylamine base and internal standard methyl *p*-tolyl sulfone were included in each of these stock solutions.

Table 20. List of operating conditions and results from the self-optimisation of the mixed variable multi-objective Heck-hydrolysis telescoped reaction. The first 25 experiments were completed as a LHC design. Pareto optimal points are highlighted in green.

Entry	t_{res}/min	Equiv. 4.6	F_vA:F_vR1	Temp/°C	Ligand	STY_{Diox} /kg m^{-3} h^{-1}	Overall Yield/%
1	9.7	1.6	1.3	85	dppm	0.60	1.73
2	15.7	3.0	0.6	119	dppm	0.43	7.73
3	1.8	2.5	0.9	142	dppm	2.07	4.44
4	19.5	1.8	1.1	174	dppm	0.59	3.53
5	5.1	1.3	1.4	184	dppm	1.63	2.67

6	9.7	1.6	1.3	85	dppe	0.01	17.10
7	15.7	3.0	0.6	119	dppe	3.28	42.17
8	1.8	2.5	0.9	142	dppe	3.94	26.41
9	19.5	1.8	1.1	174	dppe	9.21	51.98
10	5.1	1.3	1.4	184	dppe	27.83	28.26
11	9.7	1.6	1.3	85	dppp	0.24	32.90
12	15.7	3.0	0.6	119	dppp	8.09	61.18
13	1.8	2.5	0.9	142	dppp	5.96	35.42
14	19.5	1.8	1.1	174	dppp	12.81	55.85
15	5.1	1.3	1.4	184	dppp	40.48	32.39
16	9.7	1.6	1.3	85	TPP	0.16	18.08
17	15.7	3.0	0.6	119	TPP	0.72	6.17
18	1.8	2.5	0.9	142	TPP	10.82	6.30
19	19.5	1.8	1.1	174	TPP	1.24	5.87
20	5.1	1.3	1.4	184	TPP	3.48	3.74
21	9.7	1.6	1.3	85	DavePhos	0.58	10.98
22	15.7	3.0	0.6	119	DavePhos	0.96	8.87
23	1.8	2.5	0.9	142	DavePhos	4.41	4.15
24	19.5	1.8	1.1	174	DavePhos	0.93	4.57
25	5.1	1.3	1.4	184	DavePhos	2.97	2.51
26	13.3	2.6	0.5	187	dppp	1.65	27.26
27	20.0	2.1	0.8	117	dppp	2.40	69.71
28	20.0	2.1	0.7	120	dppp	2.39	65.47
29	18.4	1.1	1.0	164	dppp	1.80	54.82
30	20.0	3.0	0.8	114	dppp	2.49	73.46
31	16.0	3.0	1.0	195	dppp	1.88	50.06

32	20.0	3.0	0.7	81	dppp	0.80	22.51
33	19.5	3.0	0.9	122	dppp	2.14	67.16
34	20.0	2.6	0.5	119	dppp	2.11	52.68
35	15.8	2.5	0.8	117	dppp	2.47	56.68
36	20.0	3.0	0.9	127	dppp	2.29	71.18
37	20.0	2.2	1.0	130	dppp	2.51	83.20
38	18.1	3.0	1.0	139	dppp	2.97	88.77
39	20.0	3.0	0.9	149	dppp	2.61	81.24
40	19.9	2.9	0.6	140	dppp	2.81	74.71
41	20.0	3.0	0.6	132	dppp	3.15	82.94
42	10.0	1.5	1.0	144	dppp	3.32	54.85
43	20.0	3.0	0.9	139	dppp	2.38	74.49
44	17.4	2.9	0.8	155	dppp	1.71	43.71
45	17.4	3.0	0.6	133	dppp	3.02	69.49
46	18.9	2.4	0.6	125	dppp	2.79	70.16
47	20.0	3.0	0.6	125	dppp	2.86	76.28
48	18.3	2.0	0.6	123	dppp	2.68	64.78
49	14.5	1.5	0.6	195	dppe	2.18	41.38
50	14.6	1.7	1.0	191	dppp	2.40	58.03
51	6.9	3.0	0.5	118	dppp	1.38	11.88
52	17.9	3.0	0.7	126	dppp	2.91	71.66
53	20.0	3.0	0.5	156	dppe	2.47	61.32
54	6.3	3.0	1.0	175	dppe	4.06	42.34
55	7.0	2.1	1.0	175	dppp	5.58	64.52
56	3.1	1.0	1.0	200	dppp	6.75	34.78
57	3.5	1.5	0.9	198	dppe	4.77	27.02

58	20.0	3.0	1.0	199	dppp	2.06	68.23
59	1.0	1.0	1.0	169	dppp	18.05	29.90
60	2.4	3.0	1.0	166	dppp	13.99	55.46
61	1.0	2.0	1.0	177	dppp	26.47	43.85
62	2.9	3.0	1.0	177	dppp	6.01	28.93
63	20.0	3.0	1.0	146	dppp	1.72	56.91
64	17.9	1.9	0.7	128	dppp	2.26	56.43
65	20.0	2.5	1.0	166	dppe	0.64	21.16
66	16.8	3.0	1.0	133	dppp	2.02	56.32

Table 21. List of Pareto front experiments in order of optimum STY_{Diox} to optimum overall yield.

Entry	**t_{res}/min**	**Equiv. 4.6**	**F_vA:F_vR1**	**Temp/°C**	**Ligand**	**STY_{Diox} /kg m^{-3} h^{-1}**	**Overall Yield/%**
15	5.1	1.3	1.4	184	dppp	40.48	32.39
61	1.0	2.0	1.0	177	dppp	26.47	43.85
60	2.4	3.0	1.0	166	dppp	13.99	55.46
14	19.5	1.8	1.1	174	dppp	12.81	55.85
12	15.7	3.0	0.6	119	dppp	8.09	61.18
55	7.0	2.1	1.0	175	dppp	5.58	64.52
41	20.0	3.0	0.6	132	dppp	3.15	82.94
38	18.1	3.0	1.0	139	dppp	2.97	88.77

References

(1) Bogdan, A. R.; Dombrowski, A. W. Emerging Trends in Flow Chemistry and Applications to the Pharmaceutical Industry. *Journal of Medicinal Chemistry*. 2019, pp 6422–6468. https://doi.org/10.1021/acs.jmedchem.8b01760.

(2) Hart, T.; Schultz, V.; Thomas, D. A.; Kulesza, T.; Jensen, K. F. Development of a Versatile Modular Flow Chemistry Benchtop System. *Org. Process Res. Dev.* **2020**. https://doi.org/10.1021/acs.oprd.0c00160.

(3) Lummiss, J. A. M.; Morse, P. D.; Beingessner, R. L.; Jamison, T. F. Towards More Efficient, Greener Syntheses through Flow Chemistry. *Chem. Rec.* **2017**, *17* (7), 667–680. https://doi.org/10.1002/tcr.201600139.

(4) Beg, S. A.; Hassan, M. M. Effects of Inhibitors on Nitrification in a Packed-Bed Biological Flow Reactor. *Water Res.* **1987**, *21* (2), 191–198. https://doi.org/10.1016/0043-1354(87)90048-0.

(5) Pomberger, A.; Mo, Y.; Nandiwale, K. Y.; Schultz, V. L.; Duvadie, R.; Robinson, R. I.; Altinoglu, E. I.; Jensen, K. F. A Continuous Stirred-Tank Reactor (CSTR) Cascade for Handling Solid-Containing Photochemical Reactions. *Org. Process Res. Dev.* **2019**, *23*, acs.oprd.9b00378. https://doi.org/10.1021/acs.oprd.9b00378.

(6) Manson, J.; Clayton, A.; Gonzalez Niño, C.; Labes, R.; Chamberlain, T.; Blacker, A.; Kapur, N.; Bourne, R. A Hybridised Optimisation of an Automated Photochemical Continuous Flow Reactor. *Chimia (Aarau).* **2019**, *73* (10), 817–822. https://doi.org/10.2533/chimia.2019.817.

(7) Kim, H.; Nagaki, A.; Yoshida, J. I. A Flow-Microreactor Approach to Protecting-Group-Free Synthesis Using Organolithium Compounds. *Nat. Commun.* **2011**, *2* (1). https://doi.org/10.1038/ncomms1264.

(8) Movsisyan, M.; Delbeke, E. I. P.; Berton, J. K. E. T.; Battilocchio, C.; Ley, S. V.; Stevens, C. V. Taming Hazardous Chemistry by Continuous Flow Technology. *Chem. Soc. Rev.* **2016**, *45* (18), 4892–4928.

https://doi.org/10.1039/c5cs00902b.

(9) Plutschack, M. B.; Pieber, B.; Gilmore, K.; Seeberger, P. H. The Hitchhiker's Guide to Flow Chemistry Â¥. *Chem. Rev.* **2017**, *117* (18), 11796–11893. https://doi.org/10.1021/acs.chemrev.7b00183.

(10) Hessel, V. Process Windows - Gate to Maximizing Process Intensification via Flow Chemistry. *Chem. Eng. Technol.* **2009**, *32* (11), 1655–1681. https://doi.org/10.1002/ceat.200900474.

(11) Rott, N. *Note on the History of the Reynolds Number*; 1990; Vol. 22.

(12) Hartman, R. L.; McMullen, J. P.; Jensen, K. F. Deciding Whether to Go with the Flow: Evaluating the Merits of Flow Reactors for Synthesis. *Angew. Chemie - Int. Ed.* **2011**, *50* (33), 7502–7519. https://doi.org/10.1002/anie.201004637.

(13) Lamberto, D. J.; Alvarez, M. M.; Muzzio, F. J. Experimental and Computational Investigation of the Laminar Flow Structure in a Stirred Tank. *Chem. Eng. Sci.* **1999**, *54* (7), 919–942. https://doi.org/10.1016/S0009-2509(98)00275-9.

(14) Nagy, K. D.; Shen, B.; Jamison, T. F.; Jensen, K. F. Mixing and Dispersion in Small-Scale Flow Systems. *Org. Process Res. Dev.* **2012**, *16* (5), 976–981. https://doi.org/10.1021/op200349f.

(15) Jensen, K. F. Flow Chemistry—Microreaction TechnologyComes of Age. *AIChE J.* **2017**, *63* (3), 858–869. https://doi.org/10.1002/aic.

(16) Fernández-Maza, C.; Fallanza, M.; Gómez-Coma, L.; Ortiz, I. Performance of Continuous-Flow Micro-Reactors with Curved Geometries. Experimental and Numerical Analysis. *Chem. Eng. J.* **2022**, *437* (January). https://doi.org/10.1016/j.cej.2022.135192.

(17) Wartha, E. M.; Bösenhofer, M.; Harasek, M. Characteristic Chemical Time Scales for Reactive Flow Modeling. *Combust. Sci. Technol.* **2021**, *193* (16), 2807–2832. https://doi.org/10.1080/00102202.2020.1760257.

(18) Otálvaro-Marín, H. L.; Machuca-Martínez, F. Sizing of Reactors by Charts of Damköhler's Number for Solutions of Dimensionless Design Equations. *Heliyon* **2020**, *6* (11).

https://doi.org/10.1016/j.heliyon.2020.e05386.

(19) Fukuyama, T.; Chiba, H.; Kuroda, H.; Takigawa, T.; Kayano, A.; Tagami, K. Application of Continuous Flow for DIBAL-H Reduction and n-BuLi Mediated Coupling Reaction in the Synthesis of Eribulin Mesylate. *Org. Process Res. Dev.* **2016**, *20* (2), 503–509. https://doi.org/10.1021/acs.oprd.5b00353.

(20) Clayton, A. D.; Pyzer-Knapp, E. O.; Purdie, M.; Jones, M. F.; Barthelme, A.; Pavey, J.; Kapur, N.; Chamberlain, T. W.; Blacker, A. J.; Bourne, R. A. Bayesian Self-Optimization for Telescoped Continuous Flow Synthesis. *Angew. Chemie* **2023**, *135* (3). https://doi.org/10.1002/ange.202214511.

(21) Jiao, J.; Nie, W.; Yu, T.; Yang, F.; Zhang, Q.; Aihemaiti, F.; Yang, T.; Liu, X.; Wang, J.; Li, P. Multi-Step Continuous-Flow Organic Synthesis: Opportunities and Challenges. *Chem. - A Eur. J.* **2021**, *27* (15), 4817–4838. https://doi.org/10.1002/chem.202004477.

(22) Domokos, A.; Nagy, B.; Szilágyi, B.; Marosi, G.; Nagy, Z. K. Integrated Continuous Pharmaceutical Technologies - A Review. *Org. Process Res. Dev.* **2021**, *25* (4), 721–739. https://doi.org/10.1021/acs.oprd.0c00504.

(23) Murray, P. R. D.; Browne, D. L.; Pastre, J. C.; Butters, C.; Guthrie, D.; Ley, S. V. Continuous Flow-Processing of Organometallic Reagents Using an Advanced Peristaltic Pumping System and the Telescoped Flow Synthesis of (E/Z)-Tamoxifen. *Org. Process Res. Dev.* **2013**, *17* (9), 1192–1208. https://doi.org/10.1021/op4001548.

(24) Houben, C.; Lapkin, A. A. Automatic Discovery and Optimization of Chemical Processes. *Curr. Opin. Chem. Eng.* **2015**, *9*, 1–7. https://doi.org/10.1016/j.coche.2015.07.001.

(25) Henson, A. B.; Gromski, P. S.; Cronin, L. Designing Algorithms to Aid Discovery by Chemical Robots. *ACS Cent. Sci.* **2018**, *4* (7), 793–804. https://doi.org/10.1021/acscentsci.8b00176.

(26) McMullen, J. P.; Jensen, K. F. An Automated Microfluidic System for Online Optimization in Chemical Synthesis. *Org. Process Res. Dev.*

2010, *14* (5), 1169–1176. https://doi.org/10.1021/op100123e.

(27) Hardwick, T.; Ahmed, N. Digitising Chemical Synthesis in Automated and Robotic Flow. *Chem. Sci.* **2020**, *11* (44), 11973–11988. https://doi.org/10.1039/d0sc04250a.

(28) Löbbecke, S.; Ferstl, W.; Panić, S.; Türcke, T. Concepts for Modularization and Automation of Microreaction Technology. *Chem. Eng. Technol.* **2005**, *28* (4), 484–493. https://doi.org/10.1002/ceat.200500005.

(29) Sans, V.; Cronin, L. Towards Dial-a-Molecule by Integrating Continuous Flow, Analytics and Self-Optimisation. *Chem. Soc. Rev.* **2016**, *45* (8), 2032–2043. https://doi.org/10.1039/c5cs00793c.

(30) Weissman, S. A.; Anderson, N. G. Design of Experiments (DoE) and Process Optimization. A Review of Recent Publications. *Org. Process Res. Dev.* **2015**, *19* (11), 1605–1633. https://doi.org/10.1021/op500169m.

(31) Zhang, L.; Mao, S. Application of Quality by Design in the Current Drug Development. *Asian J. Pharm. Sci.* **2017**, *12* (1), 1–8. https://doi.org/10.1016/j.ajps.2016.07.006.

(32) Khan, A.; Imam, S. S.; Aqil, M.; Sultana, Y.; Ali, A.; Khan, K. Design of Experiment Based Validated Stability Indicating RP-HPLC Method of Temozolomide in Bulk and Pharmaceutical Dosage Forms. *Beni-Suef Univ. J. Basic Appl. Sci.* **2016**, *5* (4), 402–408. https://doi.org/10.1016/j.bjbas.2015.11.011.

(33) Czitrom, V. One-Factor-at-a-Time Versus Designed Experiments. *Am. Stat.* **1999**, *53* (2), 126–131. https://doi.org/https://doi.org/10.1080/00031305.1999.10474445.

(34) Leardi, R. Experimental Design in Chemistry: A Tutorial. *Anal. Chim. Acta* **2009**, *652* (1–2), 161–172. https://doi.org/10.1016/j.aca.2009.06.015.

(35) Clayton, A. D.; Manson, J. A.; Taylor, C. J.; Chamberlain, T. W.; Taylor, B. A.; Clemens, G.; Bourne, R. A. Algorithms for the Self-Optimisation of Chemical Reactions. *React. Chem. Eng.* **2019**, *4* (9), 1545–1554.

https://doi.org/10.1039/c9re00209j.

(36) Ouria, A.; Toufigh, M. M. Application of Nelder-Mead Simplex Method for Unconfined Seepage Problems. *Appl. Math. Model.* **2009**, *33* (9), 3589–3598. https://doi.org/10.1016/j.apm.2008.12.001.

(37) Spendley, A. W.; Hext, G. R.; Himsworth, F. R.; Technometrics, S.; Nov, N. Sequential Application of Simplex Designs in Optimisation and Evolutionary Operation. *Technometrics* **2016**, *4* (4), 441–461.

(38) Nelder, J. A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7* (4), 308–313. https://doi.org/10.1093/comjnl/7.4.308.

(39) Klein, K.; Neira, J. Nelder-Mead Simplex Optimization Routine for Large-Scale Problems: A Distributed Memory Implementation. *Comput. Econ.* **2014**, *43* (4), 447–461. https://doi.org/10.1007/s10614-013-9377-8.

(40) McMullen, J. P.; Stone, M. T.; Buchwald, S. L.; Jensen, K. F. An Integrated Microreactor System for Self-Optimization of a Heck Reaction: From Micro-to Mesoscale Flow Systems. *Angew. Chemie - Int. Ed.* **2010**, *49* (39), 7076–7080. https://doi.org/10.1002/anie.201002590.

(41) Routh, M. W.; Swartz, P. A.; Denton, M. B. Performance of the Super Modified Simplex. *Anal. Chem.* **1977**, *49* (9), 1422–1428. https://doi.org/10.1021/ac50017a033.

(42) Kazmierczak Jr., R. F. Optimizing Complex Bioeconomic Simulations Using an Efficient Search Heuristic. *SSRN Electron. J.* **2005**, No. 704. https://doi.org/10.2139/ssrn.15071.

(43) Shuster, J. J. Design and Analysis of Experiments. *Methods Mol. Biol.* **2007**, *404*, 235–259. https://doi.org/10.2307/2985452.

(44) Bourne, R. A.; Skilton, R. A.; Parrott, A. J.; Irvine, D. J.; Poliakoff, M. Adaptive Process Optimization for Continuous Methylation of Alcohols in Supercritical Carbon Dioxide. *Org. Process Res. Dev.* **2011**, *15* (4), 932–938. https://doi.org/10.1021/op200109t.

(45) Parrott, A. J.; Bourne, R. A.; Akien, G. R.; Irvine, D. J.; Poliakoff, M. Self-Optimizing Continuous Reactions in Supercritical Carbon Dioxide. *Angew. Chemie - Int. Ed.* **2011**, *50* (16), 3788–3792. https://doi.org/10.1002/anie.201100412.

(46) Moore, J. S.; Jensen, K. F. Automated Multitrajectory Method for Reaction Optimization in a Microfluidic System Using Online IR Analysis. *Org. Process Res. Dev.* **2012**, *16* (8), 1409–1415. https://doi.org/10.1021/op300099x.

(47) Jumbam, D. N.; Skilton, R. A.; Parrott, A. J.; Bourne, R. A.; Poliakoff, M. The Effect of Self-Optimisation Targets on the Methylation of Alcohols Using Dimethyl Carbonate in Supercritical CO2. *J. Flow Chem.* **2012**, *2* (1), 24–27. https://doi.org/10.1556/jfchem.2012.00019.

(48) Skilton, R. A.; Parrott, A. J.; George, M. W.; Poliakoff, M.; Bourne, R. A. Real-Time Feedback Control Using Online Attenuated Total Reflection Fourier Transform Infrared (Atr Ft-Ir) Spectroscopy for Continuous Flow Optimization and Process Knowledge. *Appl. Spectrosc.* **2013**, *67* (10), 1127–1131. https://doi.org/10.1366/13-06999.

(49) Amara, Z.; Streng, E. S.; Skilton, R. A.; Jin, J.; George, M. W.; Poliakoff, M. Automated Serendipity with Self-Optimizing Continuous-Flow Reactors. *European J. Org. Chem.* **2015**, *2015* (28), 6141–6145. https://doi.org/10.1002/ejoc.201500980.

(50) Cortés-Borda, D.; Kutonova, K. V.; Jamet, C.; Trusova, M. E.; Zammattio, F.; Truchet, C.; Rodriguez-Zubiri, M.; Felpin, F. X. Optimizing the Heck-Matsuda Reaction in Flow with a Constraint-Adapted Direct Search Algorithm. *Org. Process Res. Dev.* **2016**, *20* (11), 1979–1987. https://doi.org/10.1021/acs.oprd.6b00310.

(51) Fitzpatrick, D. E.; Battilocchio, C.; Ley, S. V. A Novel Internet-Based Reaction Monitoring, Control and Autonomous Self-Optimization Platform for Chemical Synthesis. *Org. Process Res. Dev.* **2016**, *20* (2), 386–394. https://doi.org/10.1021/acs.oprd.5b00313.

(52) Cortés-Borda, D.; Wimmer, E.; Gouilleux, B.; Barré, E.; Oger, N.; Goulamaly, L.; Peault, L.; Charrier, B.; Truchet, C.; Giraudeau, P.;

Rodriguez-Zubiri, M.; Le Grognec, E.; Felpin, F. X. An Autonomous Self-Optimizing Flow Reactor for the Synthesis of Natural Product Carpanone. *J. Org. Chem.* **2018**, *83* (23), 14286–14289. https://doi.org/10.1021/acs.joc.8b01821.

(53) Fitzpatrick, D. E.; Maujean, T.; Evans, A. C.; Ley, S. V. Across-the-World Automated Optimization and Continuous-Flow Synthesis of Pharmaceutical Agents Operating Through a Cloud-Based Server. *Angew. Chemie - Int. Ed.* **2018**, *57* (46), 15128–15132. https://doi.org/10.1002/anie.201809080.

(54) Poscharny, K.; Fabry, D. C.; Heddrich, S.; Sugiono, E.; Liauw, M. A.; Rueping, M. Machine Assisted Reaction Optimization: A Self-Optimizing Reactor System for Continuous-Flow Photochemical Reactions. *Tetrahedron* **2018**, *74* (25), 3171–3175. https://doi.org/10.1016/j.tet.2018.04.019.

(55) Christensen, M.; Yunker, L. P. E.; Adedeji, F.; Häse, F.; Roch, L. M.; Gensch, T.; dos Passos Gomes, G.; Zepel, T.; Sigman, M. S.; Aspuru-Guzik, A.; Hein, J. E. Data-Science Driven Autonomous Process Optimization. *Commun. Chem.* **2021**, *4* (1), 1–12. https://doi.org/10.1038/s42004-021-00550-x.

(56) Huyer, W.; Neumaier, A. SNOBFIT - Stable Noisy Optimization by Branch and Fit. *ACM Trans. Math. Softw.* **2008**, *35* (2), 1–25. https://doi.org/10.1145/1377612.1377613.

(57) Queipo, N. V.; Haftka, R. T.; Shyy, W.; Goel, T.; Vaidyanathan, R.; Kevin Tucker, P. Surrogate-Based Analysis and Optimization. *Prog. Aerosp. Sci.* **2005**, *41* (1), 1–28. https://doi.org/10.1016/j.paerosci.2005.02.001.

(58) Krishnadasan, S.; Brown, R. J. C.; DeMello, A. J.; DeMello, J. C. Intelligent Routes to the Controlled Synthesis of Nanoparticles. *Lab Chip* **2007**, *7* (11), 1434–1441. https://doi.org/10.1039/b711412e.

(59) Holmes, N.; Akien, G. R.; Savage, R. J. D.; Stanetty, C.; Baxendale, I. R.; Blacker, A. J.; Taylor, B. A.; Woodward, R. L.; Meadows, R. E.; Bourne, R. A. Online Quantitative Mass Spectrometry for the Rapid Adaptive Optimisation of Automated Flow Reactors. *React. Chem. Eng.*

2016, *1* (1), 96–100. https://doi.org/10.1039/c5re00083a.

(60) Cherkasov, N.; Bai, Y.; Expósito, A. J.; Rebrov, E. V. OpenFlowChem-a Platform for Quick, Robust and Flexible Automation and Self-Optimisation of Flow Chemistry. *React. Chem. Eng.* **2018**, *3* (5), 769–780. https://doi.org/10.1039/c8re00046h.

(61) Rios, L. M.; Sahinidis, N. V. Derivative-Free Optimization: A Review of Algorithms and Comparison of Software Implementations. *J. Glob. Optim.* **2013**, *56* (3), 1247–1293. https://doi.org/10.1007/s10898-012-9951-y.

(62) Brochu, E.; Cora, V. M.; de Freitas, N. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. **2010**.

(63) Manson, J. A.; Chamberlain, T. W.; Bourne, R. A. MVMOO: Mixed Variable Multi-Objective Optimisation. *J. Glob. Optim.* **2021**, *80* (4), 865–886. https://doi.org/10.1007/s10898-021-01052-9.

(64) Clayton, A. D. Self-Optimising Continuous Flow Systems. **2020**, No. February.

(65) Frazier, P. I. A Tutorial on Bayesian Optimization. **2018**, No. Section 5, 1–22.

(66) Schweidtmann, A. M.; Clayton, A. D.; Holmes, N.; Bradford, E.; Bourne, R. A.; Lapkin, A. A. Machine Learning Meets Continuous Flow Chemistry: Automated Optimization towards the Pareto Front of Multiple Objectives. *Chem. Eng. J.* **2018**, *352*, 277–282. https://doi.org/10.1016/j.cej.2018.07.031.

(67) Emmerich, M. T. M.; Deutz, A. H. A Tutorial on Multiobjective Optimization: Fundamentals and Evolutionary Methods. *Nat. Comput.* **2018**, *17* (3), 585–609. https://doi.org/10.1007/s11047-018-9685-y.

(68) Müller, P.; Clayton, A. D.; Manson, J.; Riley, S.; May, O. S.; Govan, N.; Notman, S.; Ley, S. V.; Chamberlain, T. W.; Bourne, R. A. Automated Multi-Objective Reaction Optimisation: Which Algorithm Should I Use? *React. Chem. Eng.* **2022**. https://doi.org/10.1039/d1re00549a.

(69) Cao, Y.; Smucker, B. J.; Robinson, T. J. On Using the Hypervolume Indicator to Compare Pareto Fronts: Applications to Multi-Criteria Optimal Experimental Design. *J. Stat. Plan. Inference* **2015**, *160*, 60–74. https://doi.org/10.1016/j.jspi.2014.12.004.

(70) Tai, X. Y.; Ocone, R.; Christie, S. D. R.; Xuan, J. Multi-Objective Optimisation with Hybrid Machine Learning Strategy for Complex Catalytic Processes. *Energy AI* **2022**, *7* (December 2021), 100134. https://doi.org/10.1016/j.egyai.2021.100134.

(71) Shen, F.; Wang, M.; Huang, L.; Qian, F. Exergy Analysis and Multi-Objective Optimisation for Energy System: A Case Study of a Separation Process in Ethylene Manufacturing. *J. Ind. Eng. Chem.* **2021**, *93*, 394–406. https://doi.org/10.1016/j.jiec.2020.10.018.

(72) Clayton, A. D.; Schweidtmann, A. M.; Clemens, G.; Manson, J. A.; Taylor, C. J.; Niño, C. G.; Chamberlain, T. W.; Kapur, N.; Blacker, A. J.; Lapkin, A. A.; Bourne, R. A. Automated Self-Optimisation of Multi-Step Reaction and Separation Processes Using Machine Learning. *Chem. Eng. J.* **2020**, *384* (October 2019), 123340. https://doi.org/10.1016/j.cej.2019.123340.

(73) Perera, D.; Tucker, J. W.; Brahmbhatt, S.; Helal, C. J.; Chong, A.; Farrell, W.; Richardson, P.; Sach, N. W. A Platform for Automated Nanomole-Scale Reaction Screening and Micromole-Scale Synthesis in Flow. *Science (80-.).* **2018**, *359* (6374), 429–434. https://doi.org/10.1126/science.aap9112.

(74) Baumgartner, L. M.; Coley, C. W.; Reizman, B. J.; Gao, K. W.; Jensen, K. F. Optimum Catalyst Selection over Continuous and Discrete Process Variables with a Single Droplet Microfluidic Reaction Platform. *React. Chem. Eng.* **2018**, *3* (3), 301–311. https://doi.org/10.1039/c8re00032h.

(75) Amar, Y.; Schweidtmann, A. M.; Deutsch, P.; Cao, L.; Lapkin, A. Machine Learning and Molecular Descriptors Enable Rational Solvent Selection in Asymmetric Catalysis. *Chem. Sci.* **2019**, *10* (27), 6697–6706. https://doi.org/10.1039/c9sc01844a.

(76) Shultz, C. S.; Krska, S. W. Unlocking the Potential of Asymmetric Hydrogenation at Merck. *Acc. Chem. Res.* **2007**, *40* (12), 1320–1326. https://doi.org/10.1021/ar700141v.

(77) Rubin, A. E.; Tummala, S.; Both, D. A.; Wang, C.; Delaney, E. J. Emerging Technologies Supporting Chemical Process R and D and Their Increasing Impact on Productivity in the Pharmaceutical Industry. *Chem. Rev.* **2006**, *106* (7), 2794–2810. https://doi.org/10.1021/cr040674i.

(78) Davies, I. W.; Welch, C. J. Looking Forward in Pharmaceutical Process Chemistry. *Science (80-.).* **2009**, *325* (5941), 701–704. https://doi.org/10.1126/science.1174501.

(79) Hsieh, H. W.; Coley, C. W.; Baumgartner, L. M.; Jensen, K. F.; Robinson, R. I. Photoredox Iridium-Nickel Dual-Catalyzed Decarboxylative Arylation Cross-Coupling: From Batch to Continuous Flow via Self-Optimizing Segmented Flow Reactor. *Org. Process Res. Dev.* **2018**, *22* (4), 542–550. https://doi.org/10.1021/acs.oprd.8b00018.

(80) Brouzes, E.; Medkova, M.; Savenelli, N.; Marran, D.; Twardowski, M.; Hutchison, J. B.; Rothberg, J. M.; Link, D. R.; Perrimon, N.; Samuels, M. L. Droplet Microfluidic Technology for Single-Cell High-Throughput Screening. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (34), 14195–14200. https://doi.org/10.1073/pnas.0903542106.

(81) Caruthers, J. M.; Lauterbach, J. A.; Thomson, K. T.; Venkatasubramanian, V.; Snively, C. M.; Bhan, A.; Katare, S.; Oskarsdottir, G. Catalyst Design: Knowledge Extraction from High-Throughput Experimentation. *J. Catal.* **2003**, *216* (1–2), 98–109. https://doi.org/10.1016/S0021-9517(02)00036-2.

(82) Krishnadasan, S.; Yashina, A.; deMello, A. J.; deMello, J. C. *Microfluidic Reactors for Nanomaterial Synthesis*; Elsevier Inc., 2010; Vol. 38. https://doi.org/10.1016/S0065-2377(10)38004-5.

(83) Sans, V.; Porwol, L.; Dragone, V.; Cronin, L. A Self Optimizing Synthetic Organic Reactor System Using Real-Time in-Line NMR Spectroscopy. *Chem. Sci.* **2015**, *6* (2), 1258–1264.

https://doi.org/10.1039/c4sc03075c.

(84) Bédard, A.; Adamo, A.; Aroh, K. C.; Russell, M. G.; Bedermann, A. A.; Torosian, J.; Yue, B.; Jensen, K. F.; Jamison, T. F. Reconfigurable System for Automated Optimization of Diverse Chemical Reactions. **2018**, *1225* (September), 1220–1225.

(85) Holmes, N.; Akien, G. R.; Blacker, A. J.; Woodward, R. L.; Meadows, R. E.; Bourne, R. A. Self-Optimisation of the Final Stage in the Synthesis of EGFR Kinase Inhibitor AZD9291 Using an Automated Flow Reactor. *React. Chem. Eng.* **2016**, *1* (4), 366–371. https://doi.org/10.1039/c6re00059b.

(86) Walker, B. E.; Bannock, J. H.; Nightingale, A. M.; Demello, J. C. Tuning Reaction Products by Constrained Optimisation. *React. Chem. Eng.* **2017**, *2* (5), 785–798. https://doi.org/10.1039/c7re00123a.

(87) Reizman, B. J.; Wang, Y. M.; Buchwald, S. L.; Jensen, K. F. Suzuki-Miyaura Cross-Coupling Optimization Enabled by Automated Feedback. *React. Chem. Eng.* **2016**, *1* (6), 658–666. https://doi.org/10.1039/c6re00153j.

(88) Williams, W. L.; Zeng, L.; Gensch, T.; Sigman, M. S.; Doyle, A. G.; Anslyn, E. V. The Evolution of Data-Driven Modeling in Organic Chemistry. *ACS Cent. Sci.* **2021**, *7* (10), 1622–1637. https://doi.org/10.1021/acscentsci.1c00535.

(89) Feller, D. Computational Chemistry Calculations. *J. Comput. Chem.* **1996**, *17* (13), 1571–1586.

(90) Taylor, C. J.; Manson, J. A.; Clemens, G.; Taylor, B. A.; Chamberlain, T. W.; Bourne, R. A. Modern Advancements in Continuous-Flow Aided Kinetic Analysis. *React. Chem. Eng.* **2022**, *7* (5), 1037–1046. https://doi.org/10.1039/d1re00467k.

(91) Taylor, C. J. Compunetics http://www.compunetics.net/ (accessed Aug 1, 2023).

(92) Sánchez-Ramírez, E.; Huerta-Rosas, B.; Quiroz-Ramírez, J. J.; Suárez-Toriello, V. A.; Contreras-Zarazua, G.; Segovia-Hernández, J. G. Optimization-Based Framework for Modeling and Kinetic Parameter

Estimation. *Chem. Eng. Res. Des.* **2022**, *186*, 647–660. https://doi.org/10.1016/j.cherd.2022.08.040.

(93) Villaverde, A. F.; Fröhlich, F.; Weindl, D.; Hasenauer, J.; Banga, J. R. Benchmarking Optimization Methods for Parameter Estimation in Large Kinetic Models. *Bioinformatics* **2019**, *35* (5), 830–838. https://doi.org/10.1093/bioinformatics/bty736.

(94) Jiscoot, N.; Uslamin, E. A.; Pidko, E. A. Model-Based Evaluation and Data Requirements for Parallel Kinetic Experimentation and Data-Driven Reaction Identification and Optimization. *Digit. Discov.* **2023**, 994–1005. https://doi.org/10.1039/d3dd00016h.

(95) Tsu, J.; Díaz, V. H. G.; Willis, M. J. Computational Approaches to Kinetic Model Selection. *Comput. Chem. Eng.* **2019**, *121*, 618–632. https://doi.org/10.1016/j.compchemeng.2018.12.002.

(96) Willis, M. J.; Stosch, M. von. Inference of Chemical Reaction Networks Using Mixed Integer Linear Programming. *Comput. Chem. Eng.* **2016**, *90*, 31–43. https://doi.org/10.1016/j.compchemeng.2016.04.019.

(97) Taylor, C. J.; Booth, M.; Manson, J. A.; Willis, M. J.; Clemens, G.; Taylor, B. A.; Chamberlain, T. W.; Bourne, R. A. Rapid, Automated Determination of Reaction Models and Kinetic Parameters. *Chem. Eng. J.* **2021**, *413* (June 2020), 127017. https://doi.org/10.1016/j.cej.2020.127017.

(98) Steimbach, R. R.; Kollmus, P.; Santagostino, M. A Validated “Pool and Split” Approach to Screening and Optimization of Copper-Catalyzed C–N Cross-Coupling Reactions. *J. Org. Chem.* **2020**. https://doi.org/10.1021/acs.joc.0c02392.

(99) Knox, S. T.; Parkinson, S. J.; Wilding, C. Y. P.; Bourne, R. A.; Warren, N. J. Autonomous Polymer Synthesis Delivered by Multi-Objective Closed-Loop Optimisation. *Polym. Chem.* **2022**, *13* (11), 1576–1585. https://doi.org/10.1039/d2py00040g.

(100) Clayton, A. D.; Power, L. A.; Reynolds, W. R.; Ainsworth, C.; Hose, D. R. J.; Jones, M. F.; Chamberlain, T. W.; Blacker, A. J.; Bourne, R. A. Self-Optimising Reactive Extractions: Towards the Efficient

Development of Multi-Step Continuous Flow Processes. *J. Flow Chem.* **2020**, *10* (1), 199–206. https://doi.org/10.1007/s41981-020-00086-6.

(101) Sagmeister, P.; Ort, F. F.; Jusner, C. E.; Hebrault, D.; Tampone, T.; Buono, F. G.; Williams, J. D.; Kappe, C. O. Autonomous Multi-Step and Multi-Objective Optimization Facilitated by Real-Time Process Analytics. *Adv. Sci.* **2022**, *2105547*, 1–9. https://doi.org/10.1002/advs.202105547.

(102) Reizman, B. J.; Jensen, K. F. Simultaneous Solvent Screening and Reaction Optimization in Microliter Slugs. *Chem. Commun.* **2015**, *51* (68), 13290–13293. https://doi.org/10.1039/c5cc03651h.

(103) Joseph, V. R.; Ying, H. Orthogonal-Maximin Latin Hypercube Designs. *Stat. Sin.* **2008**, *18* (1), 171–186.

(104) Zhan, D.; Cheng, Y.; Liu, J. Expected Improvement Matrix-Based Infill Criteria for Expensive Multiobjective Optimization. *IEEE Trans. Evol. Comput.* **2017**, *21* (6), 956–975. https://doi.org/10.1109/TEVC.2017.2697503.

(105) Murray, P. M.; Bellany, F.; Benhamou, L.; Bučar, D. K.; Tabor, A. B.; Sheppard, T. D. The Application of Design of Experiments (DoE) Reaction Optimisation and Solvent Selection in the Development of New Synthetic Chemistry. *Org. Biomol. Chem.* **2016**, *14* (8), 2373–2384. https://doi.org/10.1039/c5ob01892g.

(106) Felton, K. C.; Rittig, J. G.; Lapkin, A. A. Summit: Benchmarking Machine Learning Methods for Reaction Optimisation. *Chemistry–Methods* **2021**, *1* (2), 116–122. https://doi.org/10.1002/cmtd.202000051.

(107) Hone, C. A.; Holmes, N.; Akien, G. R.; Bourne, R. A.; Muller, F. L. Rapid Multistep Kinetic Model Generation from Transient Flow Data. *React. Chem. Eng.* **2017**, *2* (2), 103–108. https://doi.org/10.1039/c6re00109b.

(108) Valvi, A.; Tiwari, S. Solvent-Controlled Regioselectivity in Nucleophilic Substitution Reactions of 1-X-2,4-Difluorobenzenes with Morpholine Using Deep Eutectic Solvents. *ChemistrySelect* **2021**, *6* (2), 249–254. https://doi.org/10.1002/slct.202002806.

(109) Wendt, M. D.; Kunzer, A. R. Ortho Selectivity in SNAr Substitutions of

2,4-Dihaloaromatic Compounds. Reactions with Piperidine. *Tetrahedron Lett.* **2010**, *51* (4), 641–644. https://doi.org/10.1016/j.tetlet.2009.11.087.

(110) Wendt, M. D.; Kunzer, A. R. Ortho-Selectivity in SNAr Substitutions of 2,4-Dihaloaromatic Compounds. Reactions with Anionic Nucleophiles. *Tetrahedron Lett.* **2010**, *51* (23), 3041–3044. https://doi.org/10.1016/j.tetlet.2010.03.124.

(111) Greizerstein, W.; Brieux, J. A. The Ortho: Para Ratio in the Activation of the Nucleophilic Aromatic Substitution by the Nitro Group. *J. Am. Chem. Soc.* **1962**, *84* (6), 1032–1036. https://doi.org/10.1021/ja00865a032.

(112) Snyder, L. R.; Kirkland, J. J.; Glajch, J. L. Appendix II: Properties of Solvents Used in HPLC. In *Practical HPLC Method Development*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2012; pp 721–728. https://doi.org/10.1002/9781118592014.app2.

(113) Pabsch, D.; Figiel, P.; Sadowski, G.; Held, C. Solubility of Electrolytes in Organic Solvents: Solvent-Specific Effects and Ion-Specific Effects. *J. Chem. Eng. Data* **2022**, *67* (9), 2706–2718. https://doi.org/10.1021/acs.jced.2c00203.

(114) Rayadurgam, J.; Sana, S.; Sasikumar, M.; Gu, Q. Palladium Catalyzed C-C and C-N Bond Forming Reactions: An Update on the Synthesis of Pharmaceuticals from 2015-2020. *Org. Chem. Front.* **2021**, *8* (2), 384–414. https://doi.org/10.1039/d0qo01146k.

(115) Busacca, C. A.; Fandrick, D. R.; Song, J. J.; Senanayake, C. H. The Growing Impact of Catalysis in the Pharmaceutical Industry. *Adv. Synth. Catal.* **2011**, *353* (11–12), 1825–1864. https://doi.org/10.1002/adsc.201100488.

(116) Torborg, C.; Beller, M. Recent Applications of Palladium-Catalyzed Coupling Reactions in the Pharmaceutical, Agrochemical, and Fine Chemical Industries. *Adv. Synth. Catal.* **2009**, *351* (18), 3027–3043. https://doi.org/10.1002/adsc.200900587.

(117) Jimenez-Gonzalez, C.; Ponder, C. S.; Broxterman, Q. B.; Manley, J. B. Using the Right Green Yardstick: Why Process Mass Intensity Is Used

in the Pharmaceutical Industry to Drive More Sustainable Processes. *Org. Process Res. Dev.* **2011**, *15* (4), 912–917. https://doi.org/10.1021/op200097d.

(118) Rogers, L.; Jensen, K. F. Continuous Manufacturing-the Green Chemistry Promise? *Green Chem.* **2019**, *21* (13), 3481–3498. https://doi.org/10.1039/c9gc00773c.

(119) Wiles, C.; Watts, P. Continuous Flow Reactors: A Perspective. *Green Chem.* **2012**, *14* (1), 38–54. https://doi.org/10.1039/c1gc16022b.

(120) Bennett, J. A.; Campbell, Z. S.; Abolhasani, M. Role of Continuous Flow Processes in Green Manufacturing of Pharmaceuticals and Specialty Chemicals. *Curr. Opin. Chem. Eng.* **2019**, *26*, 9–19. https://doi.org/10.1016/j.coche.2019.07.007.

(121) Ivanković, A. Review of 12 Principles of Green Chemistry in Practice. *Int. J. Sustain. Green Energy* **2017**, *6* (3), 39. https://doi.org/10.11648/j.ijrse.20170603.12.

(122) Erythropel, H. C.; Zimmerman, J. B.; De Winter, T. M.; Petitjean, L.; Melnikov, F.; Lam, C. H.; Lounsbury, A. W.; Mellor, K. E.; Janković, N. Z.; Tu, Q.; Pincus, L. N.; Falinski, M. M.; Shi, W.; Coish, P.; Plata, D. L.; Anastas, P. T. The Green ChemisTREE: 20 Years after Taking Root with the 12 Principles. *Green Chem.* **2018**, *20* (9), 1929–1961. https://doi.org/10.1039/c8gc00482j.

(123) Ley, S. V. On Being Green: Can Flow Chemistry Help? *Chem. Rec.* **2012**, *12* (4), 378–390. https://doi.org/10.1002/tcr.201100041.

(124) Baxendale, I. R.; Ley, S. V. Polymer-Supported Reagents for Multi-Step Organic Synthesis: Application to the Synthesis of Sildenafil. *Bioorg. Med. Chem. Lett.* **2000**, *10* (17), 1983–1986. https://doi.org/10.1016/s0960-894x(00)00383-8.

(125) Sheldon, R. A. Fundamentals of Green Chemistry: Efficiency in Reaction Design. *Chem. Soc. Rev.* **2012**, *41* (4), 1437–1451. https://doi.org/10.1039/c1cs15219j.

(126) Gawande, M. B.; Bonifácio, V. D. B.; Luque, R.; Branco, P. S.; Varma, R. S. Solvent-Free and Catalysts-Free Chemistry: A Benign Pathway to

Sustainability. *ChemSusChem* **2014**, *7* (1), 24–44. https://doi.org/10.1002/cssc.201300485.

(127) Anastas, P.; Eghbali, N. Green Chemistry: Principles and Practice. *Chem. Soc. Rev.* **2010**, *39* (1), 301–312. https://doi.org/10.1039/b918763b.

(128) Blacker, J.; Breen, J.; Bourne, R.; Hone, C. The Growing Impact of Continuous Flow Methods on the Twelve Principles of Green Chemistry; 2016. https://doi.org/10.1039/9781782625940-00140.

(129) Li, Z.; Wang, S.; Chin, W. S.; Achenie, L. E.; Xin, H. High-Throughput Screening of Bimetallic Catalysts Enabled by Machine Learning. *J. Mater. Chem. A* **2017**, *5* (46), 24131–24138. https://doi.org/10.1039/c7ta01812f.

(130) Trapp, O.; Weber, S. K.; Bauch, S.; Hofstadt, W. High-Throughput Screening of Catalysts by Combining Reaction and Analysis. *Angew. Chemie - Int. Ed.* **2007**, *46* (38), 7307–7310. https://doi.org/10.1002/anie.200701326.

(131) Reetz, M. T. New Methods for the High-Throughput Screening of Enantioselective Catalysts and Biocatalysts. *Angew. Chemie - Int. Ed.* **2002**, *41* (8), 1335–1338. https://doi.org/10.1002/1521-3773(20020415)41:8<1335::AID-ANIE1335>3.0.CO;2-A.

(132) Sonogashira Coupling Oxidative Addition Trends https://chem.libretexts.org/Bookshelves/Inorganic_Chemistry/Supplemental_Modules_and_Websites_(Inorganic_Chemistry)/Catalysis/Catalyst_Examples/Sonogashira_Coupling (accessed Aug 18, 2023).

(133) Chinchilla, R.; Nájera, C. The Sonogashira Reaction: A Booming Methodology in Synthetic Organic Chemistry. *Chem. Rev.* **2007**, *107* (3), 874–922. https://doi.org/10.1021/cr050992x.

(134) Liori, A. A.; Stamatopoulos, I. K.; Papastavrou, A. T.; Pinaka, A.; Vougioukalakis, G. C. A Sustainable, User-Friendly Protocol for the Pd-Free Sonogashira Coupling Reaction. *European J. Org. Chem.* **2018**, *2018* (44), 6134–6139. https://doi.org/10.1002/ejoc.201800827.

(135) Lauterbach, T.; Livendahl, M.; Rosellón, A.; Espinet, P.; Echavarren, A.

M. Unlikeliness of Pd-Free Gold(I)-Catalyzed Sonogashira Coupling Reactions. *Org. Lett.* **2010**, *12* (13), 3006–3009. https://doi.org/10.1021/ol101012n.

(136) Yu, S.; Haight, A.; Kotecki, B.; Wang, L.; Lukin, K.; Hill, D. R. Synthesis of a TRPV1 Receptor Antagonist. *J. Org. Chem.* **2009**, *74* (24), 9539–9542. https://doi.org/10.1021/jo901943s.

(137) Brito, R.; Sheth, S.; Mukherjea, D.; Rybak, L.; Ramkumar, V. TRPV1: A Potential Drug Target for Treating Various Diseases. *Cells* **2014**, *3* (2), 517–545. https://doi.org/10.3390/CELLS3020517.

(138) Immke, D. C.; Gavva, N. R. The TRPV1 Receptor and Nociception. *Semin. Cell Dev. Biol.* **2006**, *17* (5), 582–591. https://doi.org/10.1016/j.semcdb.2006.09.004.

(139) Morales-Lázaro, S. L.; Llorente, I.; Sierra-Ramírez, F.; López-Romero, A. E.; Ortíz-Rentería, M.; Serrano-Flores, B.; Simon, S. A.; Islas, L. D.; Rosenbaum, T. Inhibition of TRPV1 Channels by a Naturally Occurring Omega-9 Fatty Acid Reduces Pain and Itch. *Nat. Commun.* **2016**, *7*. https://doi.org/10.1038/ncomms13092.

(140) Constable, D. J. C.; Curzons, A. D.; Cunningham, V. L. Metrics to “green” Chemistry - Which Are the Best? *Green Chem.* **2002**, *4* (6), 521–527. https://doi.org/10.1039/b206169b.

(141) Fairlamb, I. J. S. π-Acidic Alkene Ligand Effects in Pd-Catalysed Cross-Coupling Processes: Exploiting the Interaction of Dibenzylidene Acetone (Dba) and Related Ligands with Pd(0) and Pd(Ii). *Org. Biomol. Chem.* **2008**, *6* (20), 3645–3656. https://doi.org/10.1039/b811772a.

(142) Janusson, E.; Zijlstra, H. S.; Nguyen, P. P. T.; Macgillivray, L.; Martelino, J.; McIndoe, J. S. Real-Time Analysis of Pd2(Dba)3 Activation by Phosphine Ligands. *Chem. Commun.* **2017**, *53* (5), 854–856. https://doi.org/10.1039/c6cc08824d.

(143) Hayashi, Y. Pot Economy and One-Pot Synthesis. *Chem. Sci.* **2016**, *7* (2), 866–880. https://doi.org/10.1039/c5sc02913a.

(144) Malet-Sanz, L.; Susanne, F. Continuous Flow Synthesis. a Pharma Perspective. *J. Med. Chem.* **2012**, *55* (9), 4062–4098.

https://doi.org/10.1021/jm2006029.

(145) Clayton, A. D. Recent Developments in Reactor Automation for Multistep Chemical Synthesis. *Chemistry–Methods* **2023**, *202300021*. https://doi.org/10.1002/cmtd.202300021.

(146) Jiménez-Gonzalez, C.; Curzons, A. D.; Constable, D. J. C.; Cunningham, V. L. Expanding GSK's Solvent Selection Guide - Application of Life Cycle Assessment to Enhance Solvent Selections. *Clean Technol. Environ. Policy* **2004**, *7* (1), 42–50. https://doi.org/10.1007/s10098-004-0245-z.

(147) Yoshida, J. I.; Kim, H.; Nagaki, A. Green and Sustainable Chemical Synthesis Using Flow Microreactors. *ChemSusChem* **2011**, *4* (3), 331–340. https://doi.org/10.1002/cssc.201000271.

(148) Jiménez-González, C.; Poechlauer, P.; Broxterman, Q. B.; Yang, B. S.; Am Ende, D.; Baird, J.; Bertsch, C.; Hannah, R. E.; Dell'Orco, P.; Noorman, H.; Yee, S.; Reintjens, R.; Wells, A.; Massonneau, V.; Manley, J. Key Green Engineering Research Areas for Sustainable Manufacturing: A Perspective from Pharmaceutical and Fine Chemicals Manufacturers. *Org. Process Res. Dev.* **2011**, *15* (4), 900–911. https://doi.org/10.1021/op100327d.

(149) Britton, J.; Raston, C. L. Multi-Step Continuous-Flow Synthesis. *Chem. Soc. Rev.* **2017**, *46* (5), 1250–1271. https://doi.org/10.1039/c6cs00830e.

(150) Adamo, A.; Beingessner, R. L.; Behnam, M.; Chen, J.; Jamison, T. F.; Jensen, K. F.; Monbaliu, J.-C. M.; Myerson, A. S.; Revalor, E. M.; Snead, D. R.; Stelzer, T.; Weeranoppanant, N.; Wong, S. Y.; Zhang, P. On-Demand Continuous-Flow Production of Pharmaceuticals in a Compact, Reconfigurable System. *Science (80-.).* **2016**, *352* (6281), 61–67. https://doi.org/10.1126/science.aaf1337.

(151) Baumann, M.; Rodriguez Garcia, A. M.; Baxendale, I. R. Flow Synthesis of Ethyl Isocyanoacetate Enabling the Telescoped Synthesis of 1,2,4-Triazoles and Pyrrolo-[1,2-c]Pyrimidines. *Org. Biomol. Chem.* **2015**, *13* (14), 4231–4239. https://doi.org/10.1039/c5ob00245a.

(152) Nambiar, A. M. K.; Breen, C. P.; Hart, T.; Kulesza, T.; Jamison, T. F.; Jensen, K. F. Bayesian Optimization of Computer-Proposed Multistep Synthetic Routes on an Automated Robotic Flow Platform. *ACS Cent. Sci.* **2022**, *8* (6), 825–836. https://doi.org/10.1021/acscentsci.2c00207.

(153) Dietz, T.; Klamroth, K.; Kraus, K.; Ruzika, S.; Schäfer, L. E.; Schulze, B.; Stiglmayr, M.; Wiecek, M. M. Introducing Multiobjective Complex Systems. *Eur. J. Oper. Res.* **2020**, *280* (2), 581–596. https://doi.org/10.1016/j.ejor.2019.07.027.

(154) Sagmeister, P.; Ort, F. F.; Jusner, C. E.; Hebrault, D.; Tampone, T.; Buono, F. G.; Williams, J. D.; Kappe, C. O. Autonomous Multi-Step and Multi-Objective Optimization Facilitated by Real-Time Process Analytics. *Adv. Sci.* **2022**, *9* (10), 1–9. https://doi.org/10.1002/advs.202105547.

(155) Drahl, C. In Names, History and Legacy. *Chem. Eng. News* **2010**, *88* (20), 31–33. https://doi.org/10.1021/cen-v088n020.p031.

(156) Casey, M.; Lawless, J.; Shirran, C. The Heck Reaction: Mechanistic Insights and Novel Ligands. *Polyhedron* **2000**, *19*, 517–520. https://doi.org/10.1016/S0277-5387(99)00398-8.

(157) Cabri, W.; Candiani, I. Recent Developments and New Perspectives in the Heck Reaction. *Acc. Chem. Res.* **1995**, *28* (1), 2–7. https://doi.org/10.1021/ar00049a001.

(158) Dounay, A. B.; Overman, L. E. The Asymmetric Intramolecular Mizoroki-Heck Reaction in Natural Product Total Synthesis. *Mizoroki-Heck React.* **2009**, 533–568. https://doi.org/10.1002/9780470716076.ch16.

(159) Shibasaki, M.; Vogl, E. M.; Ohshima, T. Asymmetric Heck Reaction. *Adv. Synth. Catal.* **2004**, *346* (13–15), 1533–1552. https://doi.org/10.1002/adsc.200404203.

(160) Aggarwal, V. K.; Staubitz, A. C.; Owen, M. Optimization of the Mizoroki-Heck Reaction Using Design of Experiment (DoE). *Org. Process Res. Dev.* **2006**, *10* (1), 64–69. https://doi.org/10.1021/op058013q.

(161) Ekebergh, A.; Lingblom, C.; Sandin, P.; Wenneras, C.; Martensson, J. Exploring a Cascade Heck-Suzuki Reaction Based Route to Kinase

Inhibitors Using Design of Experiments. *Org. Biomol. Chem.* **2015**, *13* (11), 3382–3392. https://doi.org/10.1039/c4ob02694b.

(162) Faheem; Karan Kumar, B.; Chandra Sekhar, K. V. G.; Chander, S.; Kunjiappan, S.; Murugesan, S. Medicinal Chemistry Perspectives of 1,2,3,4-Tetrahydroisoquinoline Analogs - Biological Activities and SAR Studies. *RSC Adv.* **2021**, *11* (20), 12254–12287. https://doi.org/10.1039/d1ra01480c.

(163) Matcha, K.; Chernichenko, K.; Jouvin, K.; Guduguntla, S. B.; Tran, D. N.; Bell, S.; Paden, W.; Figlus, M.; Muir, C.; Elliott, A.; Diaz, C. H. Heck Reaction of 2-Oxyacrylates with Aryl Bromides: A Common Route to Monoaryl Pyruvates and Ortho Ester-Protected Monoaryl Pyruvates. *J. Org. Chem.* **2022**, *87* (16), 10736–10746. https://doi.org/10.1021/acs.joc.2c00946.

(164) Amatore, C.; Godin, B.; Jutand, A.; Lemaître, F. Rate and Mechanism of the Heck Reactions of Arylpalladium Complexes Ligated by a Bidentate P,P Ligand with an Electron-Rich Alkene (Isobutyl Vinyl Ether). *Organometallics* **2007**, *26* (7), 1757–1761. https://doi.org/10.1021/om0610849.

(165) Hyder, Z.; Ruan, J.; Xiao, J. Hydrogen-Bond-Directed Catalysis: Faster, Regioselective and Cleaner Heck Arylation of Electron-Rich Olefins in Alcohols. *Chem. - A Eur. J.* **2008**, *14* (18), 5555–5566. https://doi.org/10.1002/chem.200800411.

(166) Van Leeuwen, P. W. N. M.; Kamer, P. C. J.; Reek, J. N. H.; Dierkes, P. Ligand Bite Angle Effects in Metal-Catalyzed C-C Bond Formation. *Chem. Rev.* **2000**, *100* (8), 2741–2769. https://doi.org/10.1021/cr9902704.

(167) Greizerstein, W.; Brieux, J. A. The Ortho: Para Ratio in the Activation of the Nucleophilic Aromatic Substitution by the Nitro Group. *J. Am. Chem. Soc.* **1962**, *84* (6), 1032–1036. https://doi.org/10.1021/ja00865a032.

(168) O'Brien, A. G.; Horváth, Z.; Lévesque, F.; Lee, J. W.; Seidel-Morgenstern, A.; Seeberger, P. H. Continuous Synthesis and Purification by Direct Coupling of a Flow Reactor with Simulated

Moving-Bed Chromatography. *Angew. Chemie - Int. Ed.* **2012**, *51* (28), 7028–7030. https://doi.org/10.1002/anie.201202795.